THE UNIVERSITY
of EDINBURGH

Photophysical Characterisation of Novel Fluorescent Base Analogues

Rachel Sarah Fisher

Doctor of Philosophy by Research

The University of Edinburgh

The University of Glasgow

2018

Abstract

Fluorescent nucleic acid base analogues (FBAs) are an important class of molecule used to study the structure and dynamics of DNA and RNA. These base analogues are molecules with structures that resemble one of the natural bases but which, unlike the natural bases, have high fluorescence quantum yields. 2-Aminopurine (2AP) has long been the most widely used fluorescent base analogue and is one of the few base analogues commercially available. One problem with 2AP is that it undergoes significant quenching when incorporated into DNA: the quantum yield decreases 100 fold from that of the free base, thus becoming too low for use in, for example, single molecule studies. A secondary problem is that the 305 nm absorption peak requires excitation in the UV.

A variety of new fluorescent base analogues are being produced, with a view to remedying the deficiencies of 2AP and expanding the current range of use. The first part of this thesis explores the one-photon photophysical properties of several of these novel FBAs.

The first of these novel FBAs is the 6-aza-uridine family. These compounds, analogues of uridine, have large Stokes shifts and their absorption and emission spectra are red-shifted in comparison to 2AP; their quantum yields as free bases have been shown to exceed that of 2AP and their environmental sensitivity has been demonstrated. Time-resolved measurements reported in this thesis indicate the presence of multiple emitting species. A density functional theory (DFT) study has been carried out to rationalise these emitting species as rotational isomers. Similar fluorescence lifetime measurements were made on a second class of FBAs, the quadracyclic adenine analogues, qANs; these results also indicated the presence of multiple emitting species. Experimental results show that these FBAs undergo excited-state proton transfer. The final FBA studied in this thesis is pentacyclic adenine, pA. This FBA showed some of the most promising characteristics of all the analogues investigated, such as a high quantum yield in both polar and non-polar solvents. A time-resolved investigation into pA-containing oligonucleotides indicated that in an oligonucleotide pA adopts multiple stacked conformations and its behaviour is highly sequence dependent.

Several of these aforementioned fluorescent base analogues have absorption spectra in a region that makes them accessible to two-photon (2P) excitation with a Ti:Sapphire laser. In biological systems, multiphoton excitation has several advantages over one-photon excitation. By avoiding the use of ultraviolet light there is reduced phototoxicity. Out of focus photobleaching and autofluorescence are also minimised which leads to a higher signal-to-background ratio and allows deeper tissue penetration to be achieved. Fluorescent base

analogues tend to have small Stokes shifts; this is another problem that can be overcome by using two-photon excitation. To be of potential use in multiphoton microscopy, a FBA must have a high two-photon absorption cross-section and a high two-photon brightness. Previously, the highest two- photon brightness measured for a fluorescent base analogue was less than 2 GM. Amongst the base analogues investigated here, are several that have higher two-photon brightness than ever reported for FBAs; these include pA which is shown to have the highest 2P brightness of a FBA in an oligonucleotide, 1.3 GM, and a member of the 6-aza-uridine family which as a free base has a 2P brightness of 18 GM.

Detection of individual molecules represents the ultimate level of sensitivity and enables details about a molecular system that would otherwise be concealed using conventional ensemble techniques to be revealed. With the improved 2P brightness of the molecules measured in this thesis, it has become feasible to detect single FBA molecules using 2P excitation. To maximise the chance of detection, ultrafast, shaped laser pulses have been used as the excitation source. For the first time, the signal has been high enough and the molecule of interest sufficiently photostable such that 2P fluorescence correlation spectroscopy of a fluorescent base analogue in an oligonucleotide could be measured.

In summary, this thesis reports the fluorescence lifetimes and two-photon cross-sections of a series of novel fluorescent base analogues, as well as fluorescence correlation spectroscopy measurements of the most promising candidates.

Lay Summary

DNA is one of the most important macromolecules for life, as it is the molecule that carries the genetic code in all living organisms. Although we have known about its double helix structure for over 60 years there are still gaps in our understanding of how its structure can be flexible and how it interacts with drugs and other molecules. DNA is formed of two intertwining strands which run antiparallel to each other. Each strand is made up of repeating nucleotide units which consist of a sugar-phosphate backbone and a base. There are four different bases in DNA: adenine, guanine, cytosine and thymine.

Fluorescence techniques are some of the most useful techniques to study biomolecular systems, due to their selectivity and high sensitivity. These techniques utilise fluorescent molecules, which are molecules which absorb one colour of light and give out lower energy light of a different colour. The use of fluorescence techniques in studying DNA is hindered by the fact that the natural DNA bases, are virtually non-fluorescent. In order to study DNA, we need to make it detectable in our experiments by adding a fluorescent label. We use a type of label called a fluorescent base analogue (FBA). A FBA is a fluorescent molecule of a similar size to a natural base which forms bonds in a similar way. They can replace one of the natural DNA bases without upsetting the double helix structure.

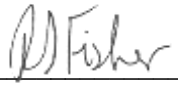
Before a fluorescent dye can be used to study DNA it first has to be characterised itself. The work in this thesis has involved the characterisation of a series of new fluorescent base analogues. This involves measurements of brightness, the colour/wavelength of the absorbed and emitted light and the time taken for absorbed light to be emitted. Part of the work in this thesis has involved two-photon excitation. Most commonly a fluorescent dye will absorb one photon (particle) of light but it is also possible for it to absorb two photons which each have half the energy required. This is advantageous as it leads to less absorbance by, and fluorescence from, the background. This results in a higher signal relative to the background and allows light to reach deeper into a sample. Measurements of two-photon cross-section have been made. These are related to the efficiency at which a molecule absorbs two-photons. Using two-photon excitation, attempts have been made to detect individual molecules of DNA each labelled with one fluorescent base analogue. This would represent the ultimate level of sensitivity.

I declare that that the work contained within this thesis is my own, except where otherwise stated and indicated with a reference.

Work from Chapter 7 of this thesis is published in the paper entitled “Pentacyclic adenine: a versatile and exceptionally bright fluorescent DNA base analogue” M. Bood, A. F. Füchtbauer, M. S. Wranne, J. J. Ro, S. Sarangamath, A. H. El-Sagheer, D. Rupert, R. S. Fisher, S. Magennis, F. Höök, A. C. Jones, T. Brown, B. H. Kim, A. Dahlen, M. Wilhelmsson and M. Grotli, *Chem. Sci.*, 2018, **9**, 3494–3502.

No part of this thesis has been submitted for any other degree or professional qualification.

Signed:



Date:

24/ 05/ 2018

Acknowledgments

Firstly I would like to thank my supervisor at the University of Edinburgh Anita Jones for all her support and guidance throughout my PhD as well as for inspiring me to start a PhD at all. Secondly, I would like to thank my supervisor at the University of Glasgow Steven Magennis for all his help and advice, without which the two-photon and single molecule measurements would not have been possible. I would also like to acknowledge David Nobis who assembled the single molecule set up used in Chapter 8 of this thesis.

Thirdly, I would also like to thank all the other people who have helped me out along the way, David Rodgers for answering my endless question when I first attempted to use Gaussian, Jochen Arlt for help with the laser and Jenny Bos for help with everything public engagement related.

Furthermore I would like to thank my high school chemistry teacher who made me realise science was interesting and made me want to study chemistry.

I would also like to thank my friends in the department. All the members of the Jones Group for their productive chats, particularly Hannah, Fi, Gareth and Grant who were so friendly and welcoming and so happy to show me how the lab (and coffee machine) worked when I first arrived. Additional thanks go to members of the Physical Chemistry section who are always willing to have a discussion and a beer. Toni, Nathan, Graham and everyone else: you have made my PhD a lot more fun! Final thanks go to Jess, science festivals would not have been the same without you- I am glad I didn't have to run a dancing light workshop by myself!

Finally, and most importantly, I would like to thank my friends, family and my sisters whom I have enthralled for the last 3 years with exciting stories about fluorescence. Most especially I would like to thank my parents who have supported me unconditionally in pretty much everything I have ever done and whom I kept forcing to read parts of this thesis.

Contents

Chapter 1: Introduction

The structure of DNA was famously solved by Watson and Crick in 1953 using the X-ray diffraction patterns measured by Rosalind Franklyn.[1] X-ray crystallography has since been used to solve thousands of structures of DNA, RNA and nucleic acid-protein complexes. A search on the Nucleic Acid Database indicates 6336 structures related to DNA (at time of writing), 5470 of those are from X-ray crystallography.[2,3] Although X-ray crystallography can give atomic-resolution structural information about DNA, it cannot capture DNA dynamics, and there is also the possibility of structural distortion due to crystal packing.[4] There are also technical limitations, such as the challenging nature of crystallising nucleic acids.[5] Consequently alternative methods are required, several of which will be summarised here.

One such alternative is nuclear magnetic resonance spectroscopy (NMR).[6] Structures determined by NMR account for most of the remaining structures in the Nucleic Acid Database. In NMR, molecules are observed in solution, closer to physiological conditions, not distorted by crystal packing. NMR can provide information about conformational mobility, the character of interactions such as hydrogen bonds and electrostatic interactions.[7] However, there are some drawbacks when it comes to investigating DNA using NMR. Proton density is lower in DNA than in most proteins, making it harder to study and, due to the linear nature, long range interactions are hard to identify using NMR techniques. NMR is mostly only useful for the study of short sequences and small DNA-protein complexes.[5] There are also technical issues with NMR, such as the need for milligram quantities of sample, often isotope-enriched samples, as well as expensive instrumentation and often time-consuming measurements.

An alternative technique that can be used to investigate the structure of DNA is electron microscopy (EM). EM has a resolution of approximately 1 nm. In fact, the first single molecule images of a DNA molecule, or any biomolecule, were performed using transmission electron microscopy.[8,9] DNA was imaged, suspended between two polystyrene beads, using a technique called shadow casting.[8] This involves deposition of a metallic film onto the sample being studied, this is done at an angle which results in parts of the sample not being coated; these areas are more electron-transparent and consequently appear as shadows. Since then, many more sophisticated measurements have been made, for example, the use of cryogenic transmission electron microscopy (cryoTEM), aided by molecular dynamics simulations, to study the structures formed by supercoiled DNA.[10] In recent years, structures have been resolved by EM in equal or greater detail than those by X-ray crystallography, the first of these being the E. coli ribosome–EF-Tu complex in which ribosomal RNA modifications could be

elucidated for the first time.[11] The main issue with using TEM to image biomolecules, aside from the highly expensive and specialised instrumentation, is that samples must be dried or frozen. This can result in the formation of non-native conformations and also does not inform on dynamics.

Another common technique which can give atomic level detail is atomic force microscopy (AFM).[12] AFM has allowed B-form DNA to be imaged in aqueous solution in enough resolution to distinguish the major and minor grooves.[13] A further commonly used technique in the study of DNA and DNA-protein complexes is small angle X-ray scattering (SAXS). SAXS does not have as high a resolution as X-ray crystallography but can be used to study disordered systems and consequently can be used to investigate the structure and conformational changes of nucleic acids in solution. It can be valuable to use these techniques in tandem with each other and with fluorescence techniques in order to maximise the information about a system; for example using SAXS, TEM and Förster resonance energy transfer to investigate the conformations of DNA origami structures.[14] Often it is also illuminating to combine experimental techniques with computation in order to gain a more complete understanding.

Fluorescence-based techniques have many advantages in the study of biomolecules.[15] They are sensitive (down to the single-molecule level), selective and they have high temporal resolution. Compared to techniques such as cryoTEM, NMR and X-ray crystallography, fluorescence-based techniques are also cheap and analysis and acquisition of data are fast and simple. Many biological processes occur on the nanosecond timescale, for example rotation of a single base about the helix axis. Fluorescence-based techniques provide the time-resolution to study these events.[16] By monitoring fluorescence properties, such as emission wavelength, quantum yield, lifetime or polarisation, information about structure, dynamics and interaction of a biomolecular system can be gained. Some biological molecules are naturally fluorescent, they contain intrinsic fluorophores.[17] The aromatic amino acids, tryptophan, tyrosine and phenylalanine, for example are found in many proteins. Intrinsic fluorophores can report on biomolecular dynamics, such a ligand binding and protein-protein interactions, without causing any perturbation to the biomolecular system. However, none of the components of DNA are intrinsically fluorescent. To circumvent this non-fluorescence, an extrinsic fluorophore can be used.[18] Fluorescent probes are fluorescent molecules that are sensitive to an aspect of their environment, such as pH, polarity or proximity to a molecule. This sensitivity manifests as a change in property such as emission wavelength or intensity.

Labelling of DNA can be either direct or indirect.[19] Direct labelling involves attaching a fluorophore covalently to the DNA backbone, non-covalently binding a dye or using an intercalating dye. Labels can be covalently bound to positions on the sugar, the phosphate group or positions on the bases. Dyes that bind non-covalently can bind to the nucleotides or the major or minor grooves. An unfortunate side effect of direct labelling is that dyes can cause perturbation to the global structure and consequently affect the biomolecular function. To minimise this perturbation fluorophores can be attached indirectly using non-fluorescent linker molecules. Often fluorescent molecules will be attached *via* linkers to the DNA backbone. The ideal molecules for attaching in this manner are relatively small, bright, photostable molecules with large Stokes shifts: often rhodamine, fluorescein or Alexa dyes are used. When attached *via* a linker, the position of the attached fluorophore can be ill-defined due to the mobility and flexibility of the linker. If the interaction of interest is between the DNA bases and, for example, an enzyme, a dye attached by a linker can be some distance from this interaction. An alternative labelling method to those previously discussed is the use of a fluorescent base analogue (FBA). FBAs are sometimes referred to as semi-intrinsic fluorophores due to fact that they are incorporated into nucleic acids in place of one of the natural bases. To be of use in this manner, a FBA must be of a similar size and shape to a native base, be non-disruptive to the helix and be highly fluorescent.

There are over a hundred FBAs in the literature, they range in size, shape and property; some are able to form Watson-crick hydrogen bonds and others are unable.[20] Their sensitivity to their surroundings varies: some are environmentally insensitive and some environmentally sensitive and able to act as probes. The FBAs of interest in this thesis are environmentally sensitive and able to form hydrogen bonds similar to those of a native base. The most widely used, environmentally sensitive FBA is currently 2-aminopurine (2AP).[21] 2AP was discovered to be fluorescent by Stryer *et al.* in 1969. It is an analogue of adenine, differing from it only by the position of an amine group.[22] It has a high quantum yield as a free base (0.68) and long fluorescence lifetime (10.6 ns), both of which are sensitive to base sequence. This sensitivity to inter-base interactions has resulted in it being used extensively in studies into DNA structure and dynamics. 2AP is by no means perfect, it undergoes significant quenching when incorporated into DNA.[15] The quantum yield decreases 100 fold from that of the free base, thus becoming too low for use in single-molecule studies.[23] A secondary disadvantage is that the 305-nm absorption peak requires excitation in the UV which can be inconvenient and costly.

A number of new fluorescent base analogues have been produced with a view to expanding the current range of use and remedying the deficiencies of 2AP.[24] Ideally, fluorescent analogues of all the natural bases would exist, each able to be incorporated in DNA with no structural perturbation. Should these FBAs also have a high brightness, be photostable and have absorption and emission in the visible region, it would become possible to detect a single molecule. Single-molecule measurements represent the ultimate level of sensitivity. Details about DNA can be obscured by conformational heterogeneity in ensemble measurements. Detecting a single molecule of DNA using a fluorescent base analogue would allow accurate knowledge of fluorophore position and orientation, a significant improvement over using an externally attached fluorophore. No currently available FBAs are suitable for such single-molecule applications, it is the aim of this thesis to investigate a series of novel FBAs in order to ascertain if any would be suitable for this task.

The structure of the thesis is laid out as follows:

Chapter 2 provides the fundamental theory relevant to fluorescence and fluorescence-based techniques utilised in this work, an introduction to computational chemistry methods, a description of the structure and dynamics of DNA, and a review of the literature on current fluorescent base analogues. The following chapter, Chapter 3, describes the main experimental techniques utilised throughout this thesis.

Chapter 4 and Chapter 5 report the investigation of photophysical properties of two families of novel fluorescent base analogues, together with related computational studies. The FBAs discussed in Chapter 4 are known as the extended-6-aza-uridines. The extended-6-aza-uridines are thymine analogues. There are six members of the extended-6-aza-uridines family, they consist of a substituted phenyl ring conjugated to a 6-azauridine core by a thiophene ring. They differ from each other in the nature of the phenyl substituent. They have red-shifted absorption and emission, compared to most fluorescent base analogues, and exhibit a sensitivity to solvent polarity.

The family of molecules discussed in Chapter 5 are known as the Quadracyclic Adenines (qANs). The qANs, consists of four members, they are based on a conjugated adenine motif, differing from each other by the position of a nitrogen. The qANs are adenine analogues with absorption and emission red-shifted in comparison to 2AP and display a sensitivity to pH.

Chapter 6 pertains to a fluorescent base analogue, pentacyclic adenine, pA. pA is also an adenine analogue and is an extended version of the qAN molecules. Here the steady-state and time-resolved fluorescence properties of pA, as a free base are presented. Due to the high

quantum yield exhibited by pA across a range of solvents, as well as a favourable absorption wavelength, pA has been incorporated into a series of oligonucleotides and the photophysical properties of pA in these oligonucleotides have been investigated.

The extended-6-azauradines, two of the qANs and pA have absorption spectra in a suitable region for two-photon excitation using a Ti:Sapphire laser. Multiphoton excitation can have several advantages over one-photon excitation. Out-of-focus photobleaching and autofluorescence are minimised which allows a higher signal-to-noise ratio and deeper tissue penetration to be achieved. Fluorescent base analogues tend to have small Stokes shifts, but this problem can be alleviated by using two-photon excitation. Chapter 7 describes the investigation of the two-photon photophysics of the extend-6-aza-uridines, the qANs and pA. (The FBAs whose properties under one-photon excitation are described in the preceding chapters.) The majority of this chapter is devoted to the measurement of two-photon cross-sections.

Based on the two-photon cross-sections reported in Chapter 7, the molecules with the most promising two-photon brightness values were investigated for their potential utility for single-molecule spectroscopy, using fluorescence correlation spectroscopy. The detection of single-molecule bursts was also attempted. These measurements were made in collaboration with David Nobis at The University of Glasgow, who built the set-up used. The results of these studies are reported in Chapter 8.

Finally, Chapter 9 summarises the main findings of this thesis, presents the overall conclusions of the collected studies and suggests future directions for research based upon these results.

1.1. References

1 J. D. Watson and F. H. C. Crick, *Nature*, 1953, 171, 737–738.

2 H. M. Berman, W. K. Olson, D. L. Beveridge, J. Westbrook, A. Gelbin, T. Demeny, S. H. Hsieh, A. R. Srinivasan and B. Schneider., *J. Biophys.*, 1992, **63**, 751–759.

3 B. C. Narayanan, J. Westbrook, S. Ghosh, A. I. Petrov, B. Sweeney, C. L. Zirbel, N. B. Leontis and H. M. Berman., *Nucleic Acids Res.*, 2013, **42**, D114-22.

4 C. Calladine, H. Drew, B. Luisi and A. Travers, *Understanding DNA: The Molecule and How It Works*, Elsevier Acedemic Press, London, third., 2004.

5 S. A. Kuznetsova and T. S. Oretskaya, *Russ. Chem. Rev.*, 2016, **85**, 445–463.

6 Z. Shajani and G. Varani, *Biopolymers*, 2007, **86**, 348–359.

7 H. M. Al-Hashimi, *J. Magn. Reson.*, 2013, **237**, 191–204.

8 C. E. Hall, *J. Biophys. Biochem*, 1956, **2**, 625–629.

9 H. Miller, Z. Zhou, J. Shepherd, A. J. M. Wollman and M. C. Leake, *Rep. Prog. Phys.*, 2018, **81**. 024601.

10 R. N. Irobalieva, J. M. Fogg, D. J. Catanese, T. Sutthibutpong, M. Chen, A. K. Barker, S. J. Ludtke, S. A. Harris, M. F. Schmid, W. Chiu and L. Zechiedrich, *Nat. Commun.*, 2015, **6**, 1–10.

11 N. Fischer, P. Neumann, A. L. Konevega, L. V. Bock, R. Ficner, M. V. Rodnina and H. Stark, *Nature*, 2015, **520**, 567–570.

12 Y. F. Dufrêne, T. Ando, R. Garcia, D. Alsteens, D. Martinez-Martin, A. Engel, C. Gerber and D. J. Müller, *Nat. Nanotechnol.*, 2017, **12**, 295–307.

13 S. Ido, K. Kimura, N. Oyabu, K. Kobayashi, M. Tsukada, K. Matsushige and H. Yamada, *ACS Nano*, 2013, **7**, 1817–1822.

14 L. K. Bruetzel, T. Gerling, S. M. Sedlak, P. U. Walker, W. Zheng, H. Dietz and J. Lipfert, *Nano Lett.*, 2016, **16**, 4871–4879.

15 J. R. Lakowicz, *Principles of Fluorescence Spectroscopy*, Springer, 3rd edn., 2006.

16 N. Nag, T. Ramreddy, M. Kombrabail, P. M. K. Mohan, J. D'souza, B. J. Rao, G. Duportail, Y. Mely and G. Krishnamoorthy, in *Reviews in Fluorescence 2006*, eds. C. D. Geddes and J. R. Lakowicz, Springer US, 1st edn., 2006, pp. 311–340.

17 J. R. Lakowicz, in *Principles of Fluorescence Spectroscopy*, Springer, 3rd edn., 2006, pp. 529–575.

18 I. Sarkar and A. K. Mishra, *Appl. Spectrosc. Rev.*, 2017, **4928**, 1–50.

19 L. J. Kricka and P. Fortina, *Clin. Chem.*, 2009, **55**, 670–683.

20 R. W. Sinkeldam, N. J. Greco and Y. Tor, *Chem. Rev.*, 2010, **110**, 2579–619.

21 A. C. Jones and R. K. Neely, *Q. Rev. Biophys.*, 2015, **2**, 1–36.

22 D. C. Ward, E. Reich and L. Stryer, *J. Biol. Chem.*, 1969, **244**, 1228–1237.

23 E. L. Rachofsky, R. Osman and J. B. A. Ross, *Biochemistry*, 2001, **40**, 946–956.

24 L. M. Wilhelmsson, *Q. Rev. Biophys.*, 2010, **43**, 159–83.

Chapter 2: Background Theory

This Chapter will discuss the principles that underlie the work of this thesis. It will begin with a description of concepts relating to fluorescence spectroscopy and computational chemistry. The chapter will then go on to outline some background information related to the structure of DNA, followed by a review of fluorescent base analogues relevant to the work of this thesis. Much of this content is derived from Lakowicz, Atkins and De Paula, and Masters and So.[1–3]

2.1 Fluorescence Spectroscopy

2.1.1 Absorption

Spectroscopy is, in essence, the observation of the interaction of electromagnetic radiation with atoms or molecules. On interaction with matter light can be scattered, reflected or absorbed. It is the absorption of either UV or visible light by molecules which will be of importance to the work in this thesis. When light is absorbed by a molecule, an electron, usually that from the ground electronic state of the molecule, is promoted to a higher energy level. The energy of the absorbed photon ($h\nu$) is equal to the energy difference between the initial (E_i) and final (E_f) molecular energy levels. This is illustrated in Figure 2-1.

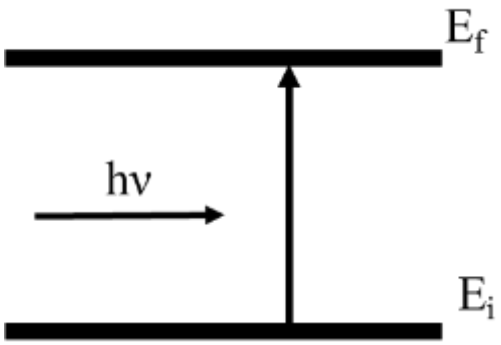


Figure 2-1. Schematic of absorption of a photon ($h\nu$) where h is Planck's constant and ν is frequency resulting in excitation of a molecule from the ground electron energy level E_i to an excited state energy level E_f.

The transition between electronic energy levels are considered to be vertical, as absorption of a photon occurs over a timescale where nuclei can be regarded as stationary. This originates from the difference in mass between nuclei and electrons: nuclei being much more massive, correspondingly have much slower motion. This allows the separation of the molecular wave function ($\phi(q, Q)$) into separate nuclear and electronic parts:

$$\phi(q, Q) \approx \psi^{elec}(q)\psi^{nuc}(Q)$$

Equation 2-1

where $\psi^{elec}(q)$ is the electronic component of the wavefunction with q representing the set of all electronic coordinates and $\psi^{nuc}(Q)$is the nuclear wavefunction with Q representing the set of all nuclear coordinates.

The absorption of UV or visible light corresponds to transitions between electronic energy levels. Each electronic state in a molecule can be considered to have a series of vibrational energy levels associated with it, electronic transitions occur from the lowest vibrational level of the ground state to vibrational levels in the excited state. The probability of a vibronic transition, and hence the spectral profile, is determined by the Franck-Condon principle. This is illustrated in Figure 2-2. The probability of a transition is dependent on the extent of the overlap between the ground and excited state vibronic wavefunctions. A greater overlap leads to a higher probability and hence a more intense transition. In Figure 2-2 the 0 – 2 transition is depicted as the most probable.

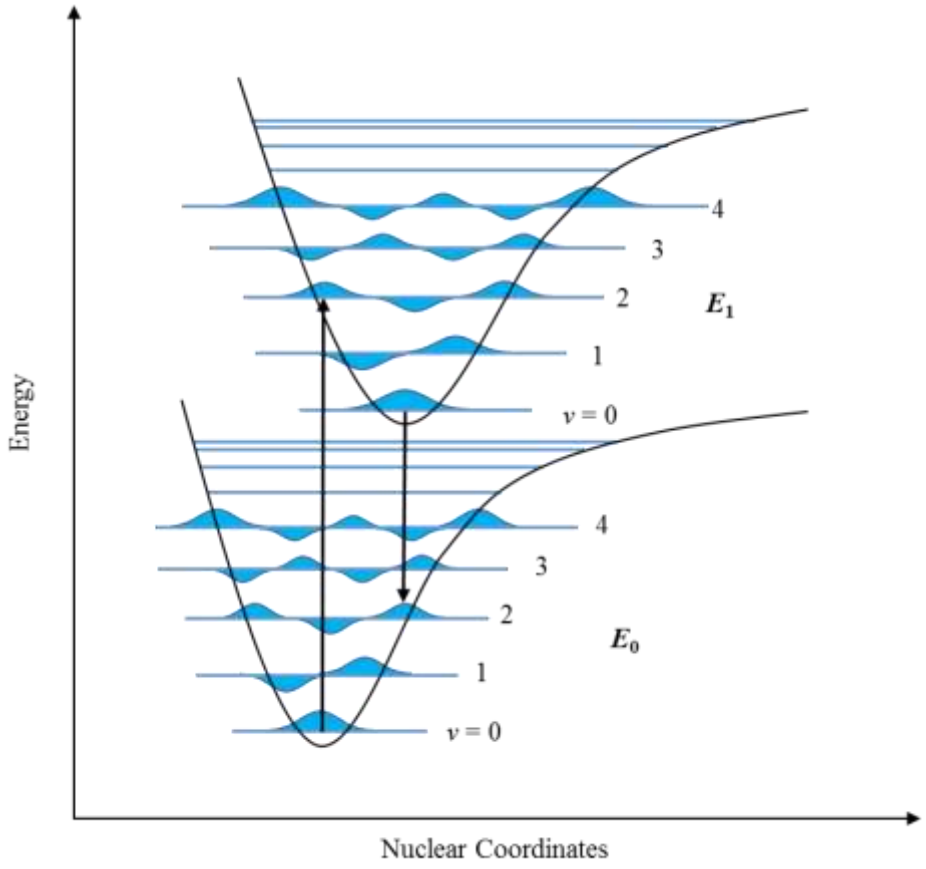


Figure 2-2. Franck-Condon energy diagram. Here a transition between S_0 ν =0 and S_1 ν =2 is illustrated. E_1 geometry is different from E_0 as represented by a shift in nuclear coordinates.

The 0-0, pure electronic, transition represents the lowest energy absorption transition, with all other transition occurring at shorter wavelengths.

Electromagnetic radiation consists of perpendicular magnetic and electric fields oscillating at some frequency. In order for electromagnetic radiation to be absorbed by a molecule there must exist, if only transiently, a dipole oscillating at the same resonant frequency. It is the coupling of the electric vector of light to the transition dipole associated with a transition that leads to absorption. The transition dipole moment (μ_{fi}) is defined by Equation 2-2:

$$\mu_{fi} = \int \Psi_f^* \hat{\mu} \Psi_i d\tau$$

Equation 2-2

where Ψ_f is the wavefunction of the final level, Ψ_i is the wavefunction of the initial level, $\hat{\mu}$ is the electric dipole operator and $d\tau$ is the integral over all dimensions of the molecule. The magnitude of the transition dipole is related to the size of the charge redistribution associated with a specific transition. The probability of an electronic transition occurring is determined by the square of the transition dipole moment. When μ_{fi} is zero, the transition is forbidden and when it is non-zero the transition is allowed. The transition dipole moment can be resolved into the three axes of a Cartesian coordinate system:

$$\mu_{fi}(x) = \Psi_f^* \hat{\mu}_x \Psi_i d\tau$$

$$\mu_{fi}(y) = \Psi_f^* \hat{\mu}_y \Psi_i d\tau$$

$$\mu_{fi}(z) = \Psi_f^* \hat{\mu}_z \Psi_i d\tau$$

Only one of these needs to be non-zero, change over the transition, in order for a transition to be symmetry- allowed.

As well as being an allowed or forbidden transition, the probability will also be dictated by the orientation of a molecule's electric dipole transition moment in relation to the plane of oscillation of the electric field of the incident light. This means than in a solution where molecules are randomly orientated, molecules with transition dipoles aligned with this plane will be preferentially excited over those orientated in a different direction.

The energy difference between electronic energy levels corresponds to light in the ultra violet or visible part of the electromagnetic spectrum. As such, absorption spectra are measured using UV-vis spectrometry. The Beer-Lambert law is used to relate the absorption of light by molecules in a solution to the concentration of that solution *c*, when the solution is in a cuvette with a path length equal to *l* cm. The Beer-Lambert law states that the number of absorbing

molecules in a sample is proportional to the amount of light absorbed and can be written in several ways:

$$A(\lambda) = 2.3 \ln \frac{I_0(\lambda)}{I(\lambda)} = N_a \sigma(\lambda) cl = \varepsilon(\lambda) cl = -\log T(\lambda)$$

Equation 2-3

Where $A(\lambda)$ is the absorbance at wavelength λ, $I_0(\lambda)$ is the intensity of incident light, $I(\lambda)$ is the intensity of transmitted light, N_a is Avogadro's number, $\sigma(\lambda)$ is absorption cross-section, $\varepsilon(\lambda)$ is the molar absorption coefficient, c is the molar concentration, l is the path-length and $T(\lambda)$ is transmission.

The strength of absorption can be given in terms of the absorption cross-section σ, with one-photon absorption cross-sections typically of the order of magnitude of 10^{-15} to 10^{-17} cm^2. More commonly for one photon processes absorption coefficient (ε) is used.

Absorbance (A) is related to the concentration of the absorbing species (c) by the molar absorption coefficient (ε) and the path length through the solution (l). The molar absorption coefficient is a measure of how strongly a molecule absorbs at a specific wavelength and is an measurable molecular property. It can be related to the transition dipole moment by a theoretical quantity; the oscillator strength (f); which is a ratio of the rate of a transition to the rate of a single electron oscillator at the same frequency:

$$f = \frac{mc_0^2 \ln 10}{N_a \pi e^2} \int \varepsilon(\bar{\nu})\, d\bar{\nu}$$

Equation 2-4

where m is the mass of an electron, e is the charge on an electron, N_a is Avogadro's number, c_0 is the speed of light and $\bar{\nu}$ is wavenumber.

Oscillator strengths can be between zero and one. A value of oscillator strength close to one corresponds to a higher probability of the transition occurring and consequently a more intense transition. An oscillator strength close to zero corresponds to a low probability electronic transition. Quantum mechanically f corresponds to the transition moment between two states over which a transition is occurring. The oscillator strength can therefore be described in terms of transition dipole moment (μ_{fi}):

$$f = \frac{8\pi^2 mc_0}{3h}\bar{v}\left|\mu_{fi}\right|^2 = 4.703 \times 10^{29}\bar{v}\left|\mu_{fi}\right|^2$$

Equation 2-5

A transition can be symmetry-forbidden; it can also be spin-forbidden. Electronic states can be described as either singlet or triplet. Singlet states are described by the symbol S_n where n= 0,1,2,3 … the number representative of the electronic state. Triplet states are described by the symbol T_n. Characterisation of a state as a singlet or triplet state is related to spin. Spin multiplicity is calculated from the total value of spin quantum numbers (S) using the formula 2S + 1. From the Pauli Exclusion Principle, each electron in an atom must have a unique set of quantum numbers, consequently, electrons in the same orbital must have opposite spins. Ground state electron spins will therefore be paired. When electron spins are paired S equals zero (+½ - ½) which results in a spin multiplicity of 1, hence this is called a singlet state. A singlet state will occur in an excited molecule when the electrons spins between a ground state and an excited electron remain paired. It is possible in the exited state for both electrons to have the same spin, without violating the Pauli Exclusion Principle, S becomes 1 (+½ +½) and spin multiplicity is 3, a triplet. Triplet-triplet or singlet-singlet transitions are allowed whereas singlet-triplet transitions are formally forbidden.

Following absorption of a photon, relaxation of the excited species can occur *via* radiative and non-radiative processes. Radiative decay processes refer to those that involve emission of light and non-radiative are those that decay without emission of light but instead by, for example, emission of heat.

2.1.2 Non-radiative Decay

As most molecules are not fluorescent, non-radiative decay is a more common fate of an excited state than radiative decay. It occurs when energy is converted into vibrations, rotations and translations of the molecule. In a fluorescent molecule, fluorescence is in competition with non-radiative decay processes such as vibrational relaxation, internal conversion and intersystem crossing. These processes are illustrated in the Jablonski diagram (Figure 2-3).

Vibrational relaxation is the fast decay of vibrational energy (under collisional conditions) back to the states accessible by thermal energy (Boltzmann distribution). Vibrational relaxation happens on a significantly shorter time scale than that of radiative decay and this has several consequences for a fluorescent molecule which will be discussed in 2.1.3.

Internal conversion (IC) and intersystem crossing (ISC) are transitions in which electronic energy is converted to vibrational energy. Internal conversion occurs between two states with

the same spin multiplicity. For example, as in Figure 2-3, between a vibrational level of S_2 and a vibrational level of equal energy in S_1. Intersystem crossing occurs between two states with different spin multiplicities, for example from S_1 to T_1. Intersystem crossing to the ground state can also occur although this is not shown in the Jablonski diagram.

The rates of IC and ISC are dependent on the extent to which overlap occurs between the wavefunction of the initial and final states. The Frank-Condon factor tells the probability of this overlap.

2.1.3 Radiative Decay

Radiative decay involves a molecule in the excited state emitting a photon to return to the ground state. There are two possible mechanisms by which this process occurs, fluorescence and phosphorescence. Both processes are illustrated in the Jablonski diagram, Figure 2-3.

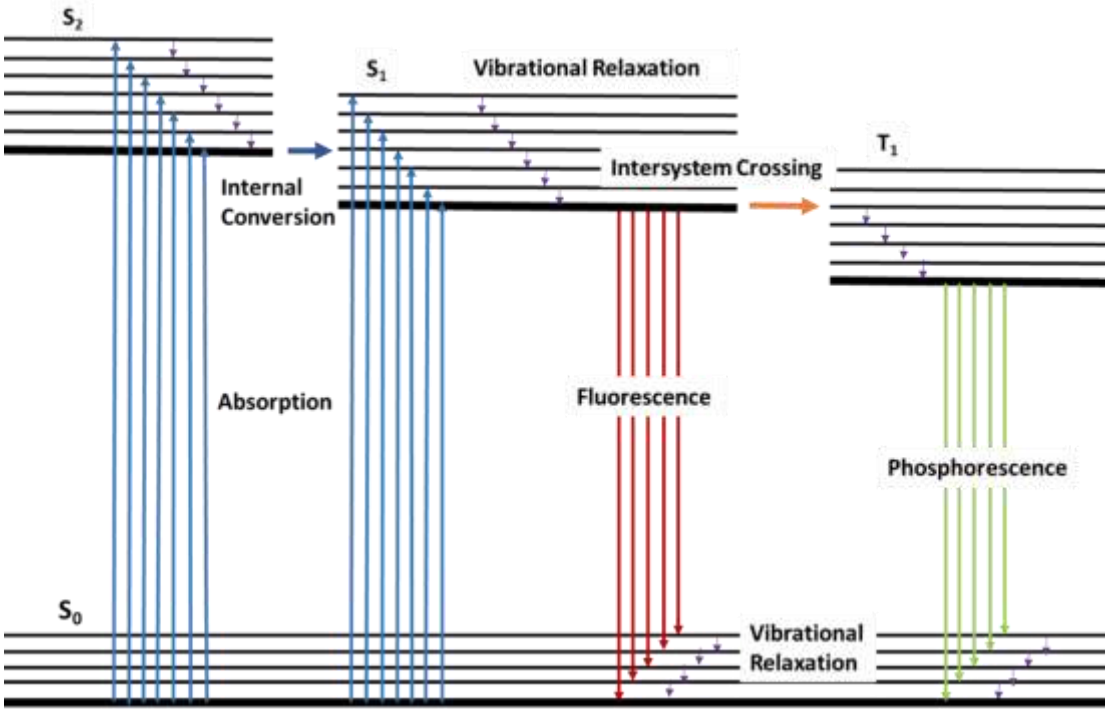


Figure 2-3. Jablonski diagram illustrating absorption, fluorescence, phosphorescence, vibrational relaxation, internal conversion and intersystem crossing.

In fluorescence, absorption from the ground state to an excited state is followed by vibrational relaxation (and internal conversion if excitation was to an excited state higher in energy than S_1) to the lowest vibrational level of S_1. Decay back to the ground state then occurs *via* emission of a photon. Vibrational relaxation and internal conversion occur rapidly. Consequently, when in solution, fluorescence emission will only take place from the lowest

excited state of the molecule. This is known as Kasha's Rule. One consequence of this is that emission wavelength is independent of the wavelength of excitation. A second consequence is that an emitted photon has a lower energy than the absorbed photon which results in a red-shifting of the emission spectra. The wavelength difference between absorption and emission is referred to as the Stokes shift. As discussed for absorption (Figure 2-2) the probability of a transition from the S_1 $\nu = 0$ to a vibrational level in the ground state is described by the Franck-Condon factor.

Fluorescence is the emission of light from excited states in which the excited state electron is spin-paired with the electron in the ground-state. Return to the ground state is spin allowed and so consequently occurs rapidly. In Figure 2-3 this is illustrated as occurring from the S_1 state to the S_0.

Phosphorescence is shown as emission of light from an excited triplet state T_1 to the S_0 ground state. As the electron in the excited orbital has the same spin orientation as the ground-state electron, transitions to the ground state are forbidden and thus emission rates are slow. Transitions are not totally forbidden, due to spin-orbit coupling. This is the name given to the interaction between the magnetic moments associated with the spin and the orbital angular momenta. Orbital angular momentum originates from an electron's motion around the nucleus. In heavier atoms with larger nuclear charge this effect is greater. This is why phosphorescence is more common in molecules containing heavy atoms.

One of the main characteristics of a fluorescent molecule is its quantum yield. Fluorscence quantum yield (ϕ) is the ratio between number of photons absorbed and number of photons emitted. It can also be represented by Equation 2-6 in which quantum yield is equal to the radiative decay rate (k_r) divided by the sum of the radiative and non-radiative decay (k_{nr}) rates.[1]

$$\phi = \frac{number\ of\ photons\ emitted}{number\ of\ photons\ absorbed} = \frac{k_r}{k_r + k_{nr}}$$

Equation 2-6

The Jablonski diagram illustrates only intermolecular processes. There are further fates of an excited state that can occur such as excited state reactions or photo-bleaching. There are also excited state processes such as intermolecular or intramolecular proton transfer that can occur and interaction with surrounding molecules that can introduce non-radiative decay pathways that lead to quenching.

The environment or solvent a fluorescent molecule is in can have a significant effect on both molecular quantum yields and spectra. Solvent polarity can have an effect on emission wavelength, as solvent can stabilise the excited state. The polarity of the excited state and the polarity of the solvent molecule will dictate the extent to which this stabilisation will occur. The Lippert-Mataga equation can be used to describe the interaction that affects the energy difference between ground and excited state, in terms of the refractive index (n) and dielectric constant (ε) of the solvent:

$$\bar{\nu}_A - \bar{\nu}_F = \frac{2}{hc}\left(\frac{\varepsilon - 1}{2\varepsilon + 1} - \frac{n^2 - 1}{2n^2 + 1}\right)\frac{(\mu_E - \mu_G)^2}{a^3} + C$$

Equation 2-7

where $\bar{\nu}_A$ is wavenumber of absorbance and $\bar{\nu}_F$ is wavenumber of emission. μ_G is the dipole moment of the ground state and μ_E is that of the excited state. The central term $\frac{\varepsilon-1}{2\varepsilon+1} - \frac{n^2-1}{2n^2+1}$ is referred to as the orientation polarizability (Δf). a is the Onsager radius and C is a constant that is equal to Stokes shift in a vacuum. The Lippert-Mataga equation only considers the effect of bulk solvation and does not account for specific solvent effects such as hydrogen bonding.

2.1.4 Fluorescence Lifetime Spectroscopy

Fluorescence decay is a first-order kinetic process with a rate constant that is the sum of the rates constants for the radiative (k_r) and all non-radiative (k_{nr}) decay processes, and can be described for a population of excited molecules M^* by Equation 2-8.[1]

$$-\frac{d[M^*]}{dt} = (k_r + k_{nr})[M^*]$$

Equation 2-8

* denotes excited state

By integrating this equation, the population of excited state molecules as a function of time can be expressed as:

$$[M^*](t) = [M^*]_0 e^{-(k_f t)}$$

Equation 2-9

where $[M^*]_0$ denotes the population at t=0 or the point of excitation and $k_f = k_r + k_{nr}$.

During fluorescence measurements the quantity measured is fluorescence intensity. Fluorescence intensity is proportional to the concentration of excited state molecules. It is related to the excited state population by:

$$I(t) = k_r[M^*](t)$$

Equation 2-10

In order to calculate the lifetime from experimentally determinable parameters, $I_0 = k_r[M^*]_0$ can be substituted into Equation 2-9 to give:

$$I(t) = I_0 e^{\frac{-t}{\tau_f}}$$

Equation 2-11

where fluorescence lifetime (τ_f) is equal to $\frac{1}{k_f}$ and I_0 is intensity at the time of the excitation pulse. Fluorescence lifetime, the time it takes for the number of excited molecules to decay to $\frac{1}{e}$ of the starting population, is the property of a molecule determined from time-resolved fluorescence measurements.

Multiple emitting species, or one emitting species in multiple environments, will result in multi-exponential decay as described in Equation 2-12.

$$I(t) = I_0 \sum_{i=1}^{n} A_i \exp(\frac{-t}{\tau_i})$$

Equation 2-12

where the pre-exponential factor A_i is the fractional contribution of each emitting species to the total excited state population when time (t) = 0, or when excitation occurs. A_i and τ_i can be used to calculate two other parameters, namely the fractional contribution of each lifetime to the steady state intensity (SS_i) and average lifetime ($<\tau_i>$). These are defined in Equation 2-13 and Equation 2-14 respectively.

$$SS_i = \frac{A_i\tau_i}{\sum A_i\tau_i}$$

Equation 2-13

$$< \tau > = \frac{\sum A_i \tau_i}{\sum A_i}$$

Equation 2-14

2.1.5 Time-correlated Single-Photon Counting

Fluorescence lifetimes in this work were measured using time-correlated single-photon counting (TCSPC). In TCSPC, pulsed laser light is used to excite molecules. Pulses are on the femtosecond timescale; orders of magnitude shorter than a typical nanosecond fluorescence lifetime. At its most basic, TCSPC acts as a stop watch. A portion of the exciting pulse is directed towards a detector and this acts as the start of the measurement; detection of the first emitted photon signals the end of the measurement. The time between the excitation pulse and the observed photon is measured and stored in a histogram. The x-axis of the histogram is the difference in time between excitation and detection and the y-axis is the number of photons detected for this time difference. This difference between the start pulse and emitted photon is collected hundreds of thousands of times, allowing a distribution of decay times to emerge.[4] In order to ensure that only one photon is emitted per excitation pulse the probability of detecting a photon from a pulse has to be kept significantly below one. If multiple photons were emitted from one pulse, as only the first photon is detected, there would be a bias towards shorter times and a true representation of the decay waveform would not be measured.

In the set-up used in the work of this thesis, the detector is a multi-channel photomultiplier (MCP). A MCP has millions of channels, each of which act as an electron multiplier. A photon that enters one of the channels hits the wall which initiates a cascade of electrons. The size of the signal produced by two identical photons, depending on the path through the channel, can vary significantly. To minimise the error associated with this variation in pulse height, a constant fraction discriminator (CFD) is used. A more detailed description of the set up used will be provided in Chapter 3, Experimental.

The shortest time resolution that can be measured by the instrument is limited by the instrument response function (IRF). Figure 4 shows an example of a typical multi-exponential decay and a typical IRF. The IRF shown in this figure was collected using excitation light scattered by colloidal silica (Ludox), a dilute scattering solution. The width of the IRF is described in terms of full width half maximum (FWHM), typically found to be ~80 ps. The width of the IRF is determined by the timing electronics and detector, the limiting factor in the set up used in this work is the transit time-spread of the MCP.

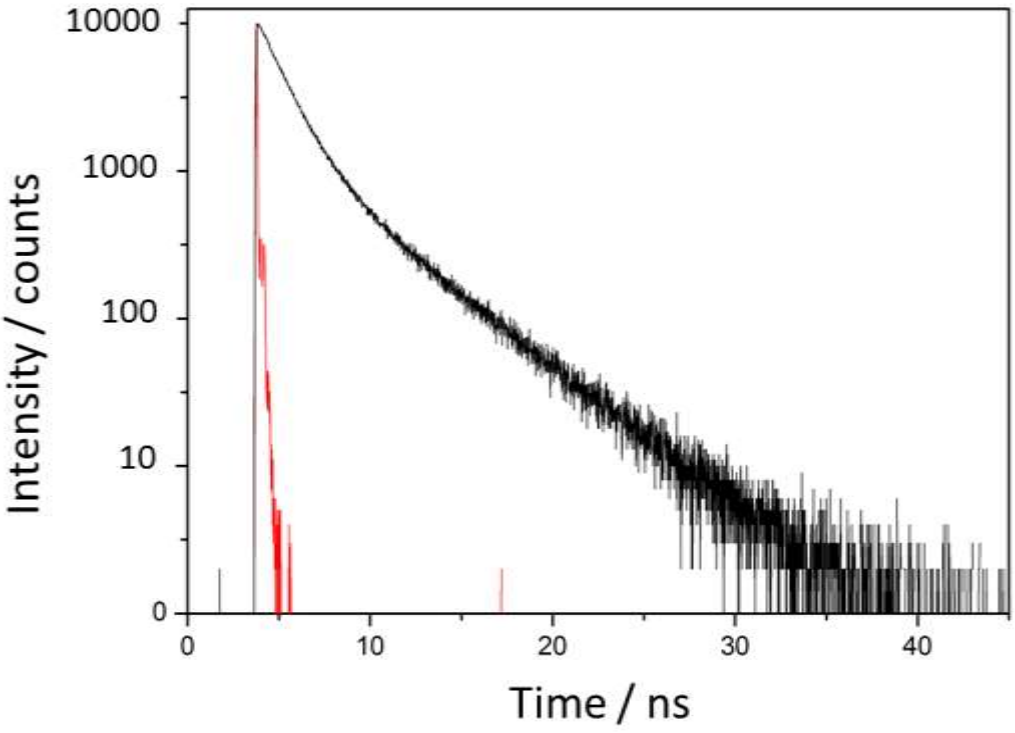


Figure 2-4. A multiexponential decay curve (black) alongside an Instrument Response Function (IRF) (red). Intensity is plotted on a logarithmic scale.

Throughout this work, an iterative reconvolution method has been for fitting decay curves. The collected fluorescence decay is a convolution of the true decay and the IRF. In iterative reconvolution, a model function, typically a multi-exponential, is convoluted with the measured IRF. The variables in this function are then varied in an iterative manner in order to optimise fit to the experimental data. The goodness of fit is determined by minimisation of the chi-squared value. The iterative reconvolution performed in this thesis was done using FAST software. An example of determination of the number of components required to obtain an adequate fit is presented in Chapter 4.

χ^2 is the value used to measure the goodness of fit of the fitted decay function and is calculated using a non-linear least squares analysis using Equation 2-15:

$$\chi^2 = \sum_{k=1}^{n} \frac{[N(t_k) - N_c(t_k)]^2}{N(t_k)}$$

Equation 2-15

where n is the number of data points, $N(t_k)$ is the measured fluorescence decay function and $N_c(t_k)$ is the calculated decay function. χ^2 is at a minimum when $N(t_k)$ and $N_c(t_k)$ are closest in value with a χ^2 of 1.0 representing a perfect fit.

The value of χ^2 for the set up used in this thesis is typically expected to be less than 1.2, based on measurements of fluorescence lifetime standards. A standard with a known single exponential decay was used in order to establish an acceptable χ^2 value, at the wavelength being used, prior to each measurement of one of the fluorescent base analogues studied in the subsequent chapters of this thesis whose lifetimes were unknown. Coumarin 153 in methanol was selected as an appropriate standard as it has been shown to have a single exponential decay with a lifetime of 4±0.2 ns when excited between 295 and 442 nm.[5]

Global fitting has been used throughout this work as this improves robustness and reliability. Global fitting involves the simultaneous analysis of multiple decays which have some parameters in common and some that differ. In this work, decays have been recorded at multiple emission wavelengths, typically three, and then simultaneously fitted assuming common lifetimes with differing populations. It is assumed that the lifetimes remain constant across different wavelengths due to the same emitting species being present. In global analysis, instead of minimising χ^2 it is the global χ^2 that is minimised.

2.2 Two-photon Absorption

Multiphoton absorption can be defined as the simultaneous absorption of two or more photons in order to excite the absorbing molecule from one electronic state to a higher energy state.[6] Three-photon, four-photon or even higher order absorption is possible, but for the purposes of this thesis only two-photon absorption will be considered. The multi-photon absorption process can be described by either semi-classical theory or quantum electrodynamical theory. The latter case, in which both the molecule and light field are described quantum mechanically, will be described here. The simple case where absorption involves two photons of equal energy is illustrated in Figure 2-5. When the two absorbed photons have the same energy this is termed degenerate multiphoton absorption. Multiphoton absorption involving photons of different energies is also possible and is termed non-degenerate absorption.

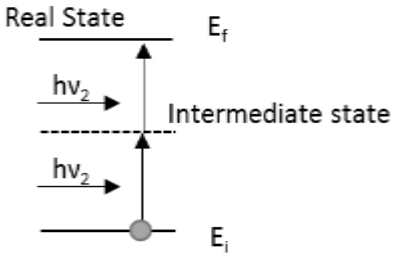


Figure 2-5. Schematic of two-photon absorption process. Solid lines represent real eigen states and dashed line represents "virtual" intermediate state.

Two-photon absorption, as illustrated by Figure 2-5, can be described as a two-step process where initially one photon is absorbed and the molecule is excited from initial state E_i to some intermediate state, represented in Figure 2-5 by a dashed line. In the second step, if a second photon arrives within the lifetime of the intermediate state, it is absorbed resulting in excitation to E_f. The amount of time spent in the intermediate state is extremely small with the two absorption occurring essentially simultaneously. In non-resonant 2P absorption, this intermediate state is considered a virtual state that is created by the absorption of the first photon. It is not a true eigenstate but a superposition of states and it can be described as a linear combination of all the real states of the molecule (ψ_i).

$$\psi_{virtual} = \sum_i c_i \psi_i$$

Equation 2-16

Where $\psi_{virtual}$ denotes the virtual state.

The individual transitions are energy non-conserving but the law of conservation of energy does hold for the total transition. The energy of the total transition is equal to the combined energy of the photons absorbed as described in Equation 2-17

$$E_f - E_i = 2hv$$

Equation 2-17

where E_f is the energy of the final state, E_i is the energy of the initial state and hv is the energy of a photon.

The attenuation of a beam of light passing through a solution as a result of two-photon absorption can be described in a similar manner to that of the Beer-Lambert law. For a one-photon process the initial intensity *I* decreases as:

$$\frac{dI(z)}{dz} = -\alpha I(z)$$

Equation 2-18

Where $I(z)$ is intensity of the incident light propagating in the z direction and α is the one-photon absorption coefficient.

The two-photon process, assuming no 1P absorption occurs and assuming the incident light has a uniform intensity which is not dependent on time, is described as:

$$\frac{dI(z)}{dz} = -\beta I^2(z)$$

Equation 2-19

Where β is the 2P absorption coefficent. Equation 2-19 illustrates that two-photon absorption is a non-linear process and proportional to the intensity squared of the incident light.

The 2P absorption coefficient β can be described in terms of the 2P cross-section, $\sigma_2(\lambda)$, and molar concentration, c using the Equation 2-20

$$\beta(\lambda) = \frac{\sigma_2(\lambda)N_0}{hv} = \frac{\sigma_2(\lambda)N_A c}{hv}$$

Equation 2-20

Where N_0 is molecular density and N_A is Avogadro's number.

This relationship is based on the assumption that the majority of molecules are in the ground state and so is only applicable to powers below that at which depletion of the ground state starts to occur. Samples in an isotropic medium will be randomly orientated, thus the value of cross-section will be an average over all orientations.

Two-photon cross-sections are much smaller that their one-photon counterparts. 1P cross-sections are on the order of ~10^{-17} cm^2 whereas typically, values for 2P cross-section are given in terms of Goeppert-Mayer (GM) units where 1GM = 1x10^{-50} cm^4 s/ photon. As these cross-sections are so small, a high intensity of light is required. A high enough intensity occurs only at the focal point of a high-powered light source, typically an ultrafast Ti-Sapphire laser, thus resulting in highly localised excitation.

Due to the small proportion of photons absorbed, directly measuring 2P absorption by quantifying the attenuation of a beam of light passing through a solution is highly complex. Where the molecule of interest is fluorescent and the quantum yield is known, a 2P fluorescence excitation method can be used. The basis for determining the two-photon cross-

section from 2P-induced fluorescence has been thoroughly described by Xu *et al.* and will be summarised here.[7]

The number of emitted photons that are detected per unit time ($F(t)$) is related to the number of photons absorbed per unit time by a 2P process ($N(t)$) by:

$$F(t) = \frac{1}{2}\phi\eta N(t)$$

Equation 2-21

where ϕ is the quantum efficiency of the dye and η is the collection efficiency of the experimental set up. $\frac{1}{2}$ represents the fact that two photons are required for each excitation event. Experimentally it is the time-averaged fluorescence photon flux $\langle F(t)\rangle$ that is actually measured.

The total number of photons absorbed per unit time depends upon the absorption cross-section (σ_2) of the sample and, as 2P absorption is a second order process, the incident intensity squared (I^2). Absorption is also a function of both the dye concentration (c) and the excitation volume (V).

$$N(t) = \int_V dV\sigma_2 c(\boldsymbol{r},t)I^2(\boldsymbol{r},t)$$

Equation 2-22

where ***r*** describes the spatial components and *t* describes the temporal components. Where the time and space dependence of the excitation intensity can be separated this equation can be written as:

$$N(t) = c\sigma_2 I_0^2(t)\int_V dV S^2(\boldsymbol{r})$$

Equation 2-23

Where $S(\boldsymbol{r})$ and $I_0(t)$ describe the spatial and temporal distribution of the incident light, ($I(\boldsymbol{r},t) = S(\boldsymbol{r})\, I_0(t)$). In writing this equation it is assumed that there is no photobleaching or ground state depletion of the sample and consequently concentration (c) is constant over both space and time.

By combining Equation 2-22 and Equation 2-23 time-averaged fluorescence photon flux $\langle F(t)\rangle$ can be described by:

$$\langle F(t)\rangle = \frac{1}{2}\phi\eta c\sigma_2\langle {I_0}^2(t)\rangle \int_V dVS^2(\boldsymbol{r})$$

Equation 2-24

Most detectors do not give a signal that is proportional to $\langle {I_0}^2(t)\rangle$ but one that is proportional to $\langle I_0(t)\rangle$. Equation 2-24 can be re-written in terms of average intensity as:

$$\langle F(t)\rangle = \frac{1}{2}g\phi\eta c\sigma_2\langle I_0(t)\rangle^2 \int_V dVS^2(\boldsymbol{r})$$

Equation 2-25

Where $g = \frac{\langle I_0^2(t)\rangle}{\langle I_0\rangle^2}$. For an ideal continuous wave laser g is equal to 1, but for a pulsed light source g is equal to $\frac{g_p}{f\tau}$ where f is repetition rate, τ is the FWHM of the excitation pulse and g_p is a quantity that is dependent on pulse shape, for a gaussian pulse g_p is found to be equal to 0.66. Although use of a continuous wave laser would simplify this equation, practically pulsed laser sources are needed in order that high powers can be achieved without damage caused by heating.

If we assume the sample is significantly thicker than the focal depth then it can be shown that

$$\int_V dVS^2(\boldsymbol{r}) \approx \frac{8n\lambda^2}{\pi^3(NA)^4}$$

Equation 2-26

Where n is the refractive index of the solution, λ is wavelength and NA is numerical apperature of the focusing lense. If the pulse intensity is the intensity averaged over the pulse duration then intensity can be related to instantaneous incident power, P by:

$$I_0(t) = \frac{\pi(NA)^2}{\lambda^2}P(t)$$

Equation 2-27

By combining Equation 2-25, Equation 2-26 and Equation 2-27 we get:

$$\langle F(t)\rangle = \frac{1}{2}\phi\eta c\sigma_2\frac{g_p}{f\tau}\frac{8n_0\langle P(t)\rangle^2}{\pi\lambda}$$

Equation 2-28

This equation describes all the experimentally measurable parameters required for determining an absolute two photon cross-section using the fluorescence excitation approach.

Absolute measurement of 2P cross-section as described here requires an accurate characterisation of the spatial and temporal properties of the excitation beam, which is a non-trivial task. By using a reference standard under the same excitation conditions as the sample of interest, the term describing the spatial distribution of light $\int_V dVS^2(\boldsymbol{r})$ and g the term describing the degree of temporal coherence will be the same and so direct measurement of these parameters is not required. The ratio of measured fluorescence signals can then be described as:

$$\frac{F^R}{F^S} = \frac{\sigma_2^R \phi^R \eta^R n^R c^R \langle P^R \rangle^2}{\sigma_2^S \phi^S \eta^S n^S c^S \langle P^S \rangle^2}$$

Equation 2-29

where ϕ is the quantum yield of fluorescence, use of this equation assumes that quantum yield is the same for 1P and 2P process. η is a term that accounts for the wavelength-dependent collection efficiency of the fluorescence. *n* is the refractive index of the solvent. *c* is the concentration. *F* is the integrated fluorescence signal from the recorded spectrum. *P* is the excitation power and the superscripts S and R refer to either the sample or reference.

In this work, the two-photon cross-sections have been measured relative to a reference using the equation:[8]

$$\sigma_2^S = \frac{\sigma_2^R \phi^R \eta^R n^S c^R F^S \langle P^R \rangle^2}{\phi^S \eta^S n^R c^S F^R \langle P^S \rangle^2}$$

Equation 2-30

References used in this thesis and verification of the accuracy of measurements made using this method will be described in detail in Chapter 7.

Equation 2-30 indicates that there is a relationship between fluorescence intensity and power squared. The square-law dependence of two-photon-excited fluorescence can be utilised to confirm that a 2P process is occurring. Deviations from this law can result for several reasons; for example, excited-state absorption, stimulated emission and excited-state saturation can all occur at high powers. Alternatively, photobleaching of the sample can also result in deviations.

So far 2P absorption has been discussed in terms of how to experimentally determine a two photon cross-section. In order to discuss the effect of local environment and molecular structural factors on 2P cross-section it is useful to consider the theoretical description. Time-dependent perturbation theory can be used to describe the 2P cross-section for randomly

orientated molecules in solution, where the two excitation photons are identical, using the equation:

$$\sigma_{2P}(\omega) = \frac{2(2\pi)^5}{15(nhc_0)^2} L^4 \left|\vec{\mu}_{if}\right|^2 \left|\Delta\vec{\mu}_{if}\right|^2 g_{i\to f}(2\omega)$$

Equation 2-31

where h is Plank's constant, c_0 is the speed of light, n is the refractive index, L is the Onsager local field factor $\frac{3n^2}{(2n^2+1)}$, $\vec{\mu}_{if}$ is the transition dipole moment vector for the transition between the initial ground state and the excited state, $\Delta\vec{\mu}_{if}$ is the change in permanent dipole and $g_{i\to f}(2\omega)$ is the line shape function. This equation describes a simplified two level system, as opposed to describing all excited states. For centrosymmetric molecules the term $\left|\Delta\vec{\mu}_{if}\right|$ is zero but will contribute for non-centrosymentric molecules.

In section 2.1.1 it was shown that the 1P cross-section is proportional only to the transition dipole moment. However, the 2P cross-section also depends on the change in permanent dipole moment on excitation. This means that, as both the ground and excited state dipole moments are important factors in the 2P absorption process, then changes in environmental factors that affect the change in dipole moment will have a larger effect on the 2P cross-section than the 1P.

This was demonstrated well for a series of proteins called the "Fruit" proteins.[9,10] The fruit proteins all contain the same chromophore but the electrostatic environment of the protein barrel in which it is located varies slightly. This has a small effect on the 1P absorption coefficient but causes a much larger change in 2P cross-section. This is because the chromophore has a difference in permanent dipole in the ground and excited state. An external electric field can induce a dipole moment which can either increase or decrease the intrinsic permanent dipole moment. The induced dipole depends on the polarizability of the ground and excited states. Changes in the protein barrel that result in change to the electrostatic environment can therefore have a significant effect on a molecule, assuming a molecule is polarizable and has a difference in polarizability between ground and excited states. Not only will the 2P cross-section be affected by the electrostatic surroundings of a protein but also by the effects of solvation. The 2P cross-section may not be at a maximum in the same solvent where 1P cross-section is at a maximum.

The first systematic study into solvent effects on 2P absorption cross-sections using a femtosecond excitation source were performed by Woo *et al.* in 2005.[11] They investigated a

series of distyrlbenzene compounds in toluene, THF, acetone, DMSO and water and found that 2P cross-section did not correlate with polarity, with the maximum 2P cross-section found in THF, a solvent of intermediate polarity. Conversely Johnston and Ogilby showed for a series of vinyl benzene derivatives that the cross-section was higher in toluene than THF.[12] Like Woo *et al.* they found that they could not discern a correlation between cross-section and solvent parameters such as polarity, illustrating that the effect seen is dependent on the molecule and the solvent. They conclude that, as the virtual state in 2P absorption is a sum over all states in the system, for different molecules different states will have a larger contribution to the linear combination than others and this difference in virtual state can explain the solvent effects. There are examples where a correlation between solvent polarity and 2P absorption cross-section is observed. 2,6-diphenyl-4-(2,4,6-triphenyl-1-pyridinio)phenolate (Reichardt's dye) shows decreasing 2P cross-section with increasing polarity, when investigated in three solvents chloroform, DMF and DMSO.[13] Theoretical calculations that considered transitions between the ground and first excited state confirmed this trend. The same group investigated 4-dimethylamino-4'-nitrostilbene (DANS) in the same three solvents, but this time observed only a small change in cross-section.[14] A theoretical study showed that for DANS the highest 2P cross-section occurs in solvents of intermediate polarity.[15]

The effect on 2P absorption of ground and excited state dipole moment is also a consideration when designing molecules with high 2P cross-sections.[16] To design molecules with high two-photon cross-sections it is often desirable to have molecular structures that encourage intramolecular charge-transfer. The most common motif uses π-electron-rich donors and π-electron-deficient acceptors conjugated together in 2D or 3D structures. Increasing the number of conjugated branches, in either 2 or 3 dimensions, can lead to an increase in 2P cross-section.

2.3 Single Molecule Detection

Detection of a single molecule represents the highest level of sensitivity and detail. There is intrinsic molecular heterogeneity when studying biomolecules which can be obscured at the ensemble level. For biomolecules such as DNA this heterogeneity can be both temporal and spatial. Biomolecules can exist in various conformations and can move between these conformations at different times. One well-studied example of the use of single molecule spectroscopy to elucidate the movements of a biomolecule that had previously been obscured by ensemble averaging, is the motion of myosin along actin.[17] Myosins are a family of molecular motors, proteins that convert ATP into mechanical energy and are involved in muscle contraction. Myosin slides along actin filaments in steps; each step is fuelled by energy from ATP hydrolysis. Ensemble techniques can give information on the average translocation

of myosin along actin. Single-molecule techniques can give information on the step-size myosin takes, the speed and frequency of the steps and also the total distance travelled.

There are multiple techniques that can give information about biomolecules at the single-molecule level but the most widely used and most useful for studying dynamics are fluorescence techniques. The first direct optical detection of a single molecule in solution was performed by Hirschfeld in 1976.[18] Using total internal reflection fluorescence microscopy (TIRF) he detected fluorescence bursts from a single globulin protein labelled with between 80 and 100 FITC fluorophores. TIRF is able to resolve single molecules due to the small excitation volume it creates. It works on the principle that, at the interface between two media of different refractive indices, an evanescent wave is generated. When excited at an angle greater than the critical angle, excitation light is reflected and the generated evanescent waves (that can only penetrate ~100 nm into the sample) selectively excite fluorophores close to the surface and not those in the bulk solution. This leads to a low background which is crucial for single-molecule detection.

The first detection of individual fluorophores in solution was not achieved until 1990 when bursts from individual Rhodamine 6G molecules were detected in water and ethanol.[19] This was achieved using a flow cytometry set up. A Rhodamine solution (100 fM) was flowed through the excitation beam and the resulting fluorescence was collected at 90 degrees to excitation. Burst of 10-15 photons per molecule were observed in water and 80-120 in ethanol. Shortly after this, the first detection of a singly labelled biomolecule, a lipid molecule labelled with a single Rhodamine molecule, was performed by Schmidt *et al*. Single-molecule detection was extended to two-photon excitation by Xu and Webb in 1995.[20] They utilised an epifluorescence set-up and were able to detect Rhodamine B with an average signal-to-background ratio (SNR) of 10 and a peak signal-to-background ratio (SBR) of 60. The advantages of two-photon excitation will be discussed in Chapters 7 and 8.

Since this proof-of-principle work carried out in the 1990s, single-molecule spectroscopy has become a more sophisticated and widely used technique.[21] Diffusional movements or movements such as that of myosin can be resolved. Protein conformational changes or interactions that result in changes in quantum yield or spectrum can be measured.[22] Multiple fluorescent probes can be used to investigate the interaction of two molecules, for example DNA and a protein, and polarization of excitation and emission can resolve rotational information.

The basic set-up required to detect single molecules in solution is as follows. A laser beam is focused into the sample using an objective with a high numerical aperture. A small diffusion-

limited excitation volume is advantageous as it results in a lower background. The duration of time spent in the focal volume and the fluorescence lifetime of the molecule determines the number of photons that can be detected. Molecules can be brought into the focal volume using flow cytometry which involves molecules being flowed into the detection volume in a controlled manner. An alternative approach is to allow molecules free in solution to diffuse into the focal volume. This is the approach used in the microscopy employed in this thesis. Fluorescence from the sample is collected using the same objective as for excitation. Due to the small focal volume, this is easier than perpendicular detection. Emission is separated from the excitation light using a dichroic mirror. Photons are then detected by a detector sensitive enough to detect single photons, commonly a photomultiplier (PMT) or Single Photon Avalanche Diode (SPAD). The detectors utilised in the single-molecule microscopy in this thesis are SPADs (MPD PDM 50c and MPD $PD-050-CTB). SPADs have high temporal resolution, <50 ps for the ones used in this work. SPADs are solid state detectors, a single photon arriving at a reverse bias p-n junction generates an electron-hole pair which leads to an avalanche, the rising edge of this avalanche pulse signals the arrival of a photon. The current is then quenched by lowering the bias voltage. In order to detect the next photon the bias voltage is restored. The main advantages of a SPAD over a PMT are efficiency and a low background count rate. SPADs have small active areas (to reduce dark count) compared to PMTs which is why they tend not to be used for bulk fluorescence spectroscopy. Details of the set-up used in this thesis will be provided in Chapter 8.

Single-molecule measurements involve the detection of fluorescence bursts emanating from a single dye molecule. To ensure that detection occurs from just one molecule, not two or more, the average number of molecules in the detection volume must be significantly lower than one. To achieve this either low concentrations or small detection volumes are required. In order to distinguish a single molecule at this level of sensitivity, background must be sufficiently low that a signal-to-noise ratio of greater than one is achieved. The main sources of background are Rayleigh scattering, Raman scattering or autofluorescence.[23]

The single molecule signal must be greater than fluctuation in signal, background and dark count in order to be detectable. Signal (S), the number of counts in a given integration time Δt is equal to:

$$S = \eta\phi\frac{\sigma}{A}\frac{P}{h\nu}\Delta t$$

Equation 2-32

where η is instrument collection efficiency, ϕ is quantum yield, σ is the absorption cross-section, A is incident beam area and $\frac{P}{hv}$ describes the number of incident photons per second. Signal to noise, assuming noise fluctuations are Poisson noise fluctuations from the background and dark counts (the number of registered counts without incident light) is given by:

$$\frac{S}{N} = \frac{\eta\phi\frac{\sigma}{A}\frac{P}{hv}\Delta t}{\sqrt{\left(\frac{\eta\phi\sigma P\Delta t}{Ahv}\right) + C_b P\Delta t + D\Delta t}}$$

Equation 2-33

Where C_b is the background count rate per watt of excitation power, and D is the number of dark counts per second. The denominator terms, as they represent intensity fluctuations not intensity, are raised to one-half power.

A multi-channel scaler (MCS) trace, with bins of 1 ms, for Rhodamine 101 (4×10^{-11} M) in water and the corresponding trace for pure water are displayed in Figure 2-6. By plotting the number of detected photons per time interval as a function of observation time, individual molecule bursts can be identified as peaks in the data. At the most basic level this indicates the presence of individual molecules.

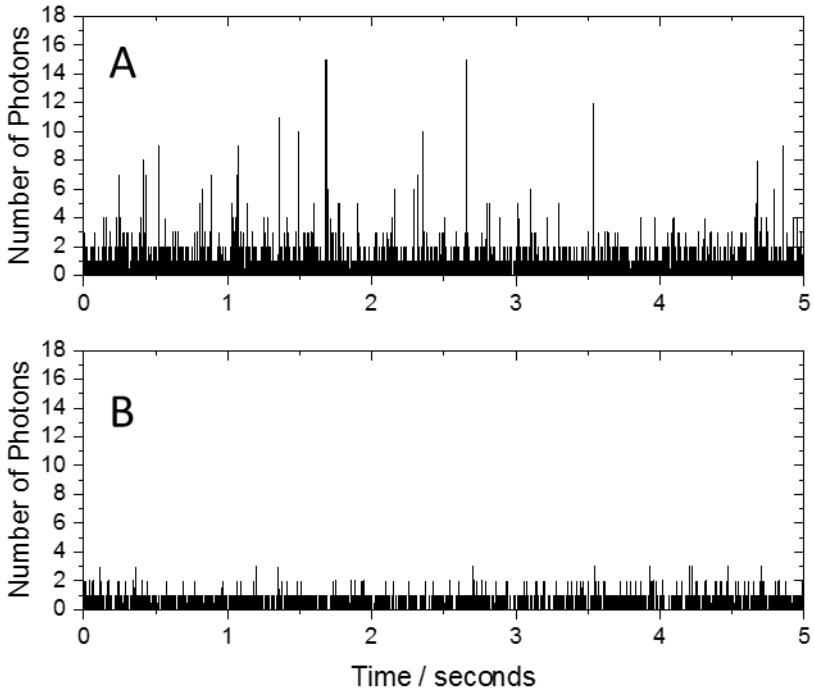


Figure 2-6. MCS trace of A) Rhodamine 101 in millipure water and B) millipure water

Analysis is required to determine if each burst is due to one molecule, or if multiple fluorophores are contributing. Single molecules can be identified by their fluorescence properties such as spectrum or lifetime. The presence and identity of an individual molecule is the simplest information that can be gathered from this type of measurement. Extrapolating this to more complex systems, information can be gathered about protein conformation, diffusion times etc.

The approach described so far has only considered the detection of single-molecule bursts. There are many techniques that can be performed at the single-molecule level that give a wide range of information about a system. Microscopy is the intuitive way to study biomolecules. Unfortunately, the spatial resolution of any optical system is diffraction-limited; due to the wave nature of light. Biological molecules such as DNA are several orders of magnitude smaller than the diffraction limit and consequently cannot be resolved through optical microscopy. If two fluorophores are closer together than ~ half the wavelength at which they are being observed they cannot be independently resolved.

The minimum distance that allows the resolution of two point source emitters (d) is defined as

$$d = \frac{\lambda}{2NA}$$

Equation 2-34

Where λ is the wavelength of light emitted and NA is the numerical aperture of the microscope objective. NA is defined as $n \sin\theta$ where n is the refractive index of the medium the lens is working in, often air, oil or as in the work presented in Chapter 8 of this thesis water (n= 1.33), and θ is the half the angular aperture.

Super-resolution imaging techniques make use of photophysical properties such as blinking or photobleaching in order to resolve beyond the diffraction limit. For example, photoactivated localization microscopy (PALM) allows resolution of fluorophores closer together than the diffraction limit by taking large numbers of images of the same protein, labelled with numerous fluorophores, with only certain fluorophores activated in each image. The activation is carried out using a low-power laser pulse. Imaging and photobleaching the fluorophores back to their non-emissive state is then performed using a high power laser beam. In each image the activated fluorophores will be separated by a great enough distance that individual spots can be identified. By fitting to, typically a Gaussian function, the centre of diffraction limited spots that correspond to individual fluorophores can be found. The images can then be overlaid to create the super resolution image. In conjunction with TIRF, PALM is able to

resolve fluorophores to within ~10 nm. STORM is a similar technique to PALM but fluorophores are switched between emitting states as opposed to switching from emissive to non-emissive states. Bleaching/binding assisted localisation microscopy (BaLM) also works on this principle but relies on the fluorophore blinking, where the molecule undergoes intersystem crossing to the triplet state. Stimulated-emission depletion microscopy works on a similar localisation of individual fluorophores principle. Once the sample has been excited, a donut-shaped beam is used to de-excite, thus the location of detected fluorophores can be attributed to the small area in the centre of the donut. These techniques have been extended into 3D, where resolution of ~20 nm (xy) and ~50 (z) has been achieved for PALM. These techniques rely on molecules remaining in the same position for the duration of the measurement. They are therefore normally performed on molecules tethered to surfaces. Super resolution microscopy techniques open the possibility of visualising the distribution of DNA inside a cell. This is challenging due to the labelling density it is possible to achieve on DNA. Nonetheless intercalating dyes, such as YOYO have been used to image DNA.[24,25] Single molecule imaging of DNA has also been used to study the localisation of DNA using 5-ethynyl-2'-deoxyuridine and an Alexa dye.

To gain detailed information about distance, single molecule Förster resonance energy transfer (FRET) can be used. FRET is based upon the high distance-dependence of the dipole-dipole interaction between two molecules. An excited donor fluorophore can non-radiatively transfer its excitation energy to an acceptor within ~ 10 nm of it. The acceptor can then emit a photon. The efficiency of this process is dependent on the distance between the donor and acceptor; approximately dependent on the inverse sixth power of the distance between the donor and acceptor pair. Single-molecule FRET has been used to study several aspects of DNA, including structural information about three- and four-way junctions. Should a fluorescent base analogue be bright enough, such that it could be detected at a single-molecule level, FRET could be one technique where this would be enormously beneficial. When a dye is attached to the DNA backbone its position and orientation are not well defined. The orientation and position of a FBA on the other hand is precisely known, which results in more accurate measurements.

2.3.1 Fluorescence Correlation Spectroscopy

Fluorescence correlation spectroscopy (FCS) is a technique based on detecting fluctuations in intensity that result from single molecules, freely diffusing in solution, passing in and out of a focal volume, typically around 10 fl. The number of molecules present at any one time is determined by Poisson statistics. In order to observe fluctuations, FCS is best performed on

very dilute solutions, so that only a small number of molecules are present in the focal volume at any time. Figure 2-7 displays some typical intensity data collected for Rhodamine B.

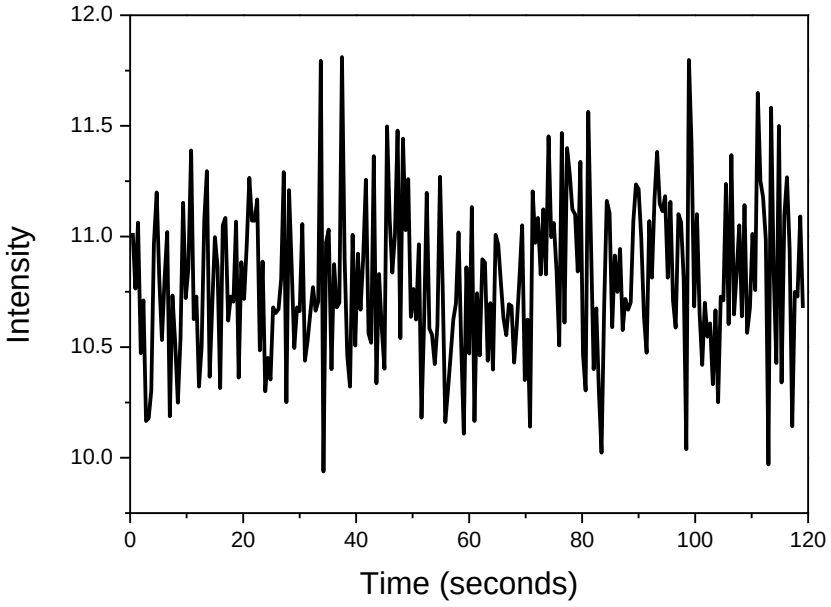


Figure 2-7. Fluorescence intensity of Rhodamine B in water collected over 120 seconds.

Fluctuations over time are calculated as:

$$\delta F(t) = < F > - F(t)$$

Equation 2-35

where $F(t)$ is fluorescence intensity at some time t and $< F >$ is the average signal. From this data a correlation curve is produced. Autocorrelation measures how the intensity after a certain time differs from the average intensity $< F >$. An autocorrelation curve results from these fluctuations and the autocorrelation function of the fluorescence fluctuations $G(\tau)$ is described by:

$$G(\tau) = \frac{< F(t)F(t+\tau) >}{< F >< F >} = 1 + \frac{< \delta F(0)\delta F(\tau) >}{< F >^2}$$

Equation 2-36

where τ is delay time between two intensity point.

From the intensity data in Figure 2-7 a correlation curve for Rhodamine B is calculated, this is shown in Figure 2-8.

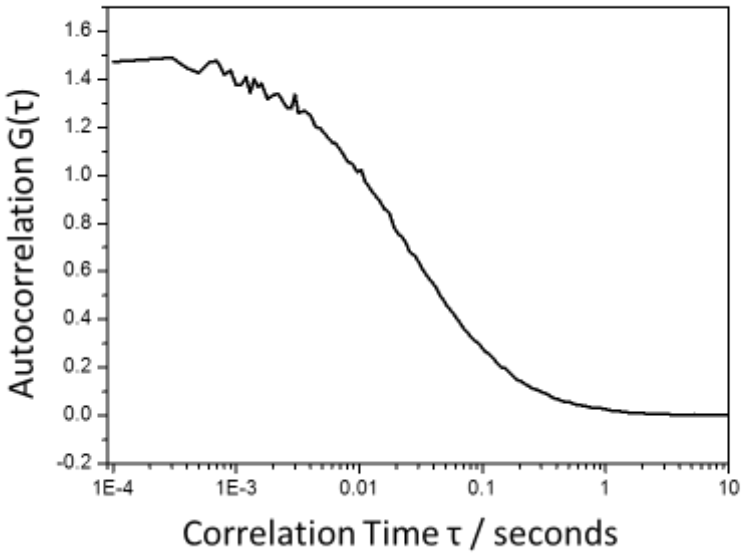


Figure 2-8. Fluorescence correlation curve resulting from the intensity data displayed in Figure 2-7.

Fitting this correlation curve can give information such as number of molecules in the focal volume, counts per molecule, diffusion coefficient and diffusion time.

The average number of molecules in the focal volume <N> can be calculated from the amplitude of the autocorrelation function (G(0)):

$$< N > = \frac{\gamma}{G(0)}$$

Equation 2-37

Where γ is a geometric factor dependent on sample volume. For a Gaussian profile γ is equal to $\frac{1}{\sqrt{8}}$.

Assuming fluctuations are caused only by diffusion, there is a single species and a Gaussian observation volume then:

$$G(\tau) = \frac{1}{[N\left(1+\frac{\tau}{\tau_D}\right)\left(1+\frac{\tau}{\omega^2\tau_D}\right)^{0.5}]}$$

Equation 2-38

where z_0 represents the axial dimension of the Gaussian excitation volume and r_0 lateral dimension with $\omega = \frac{z_0}{r_0}$ and τ_D is diffusion time, the average amount of time a molecule spends diffusing through the observation volume.

From diffusion time, assuming the focal volume is well characterised, a diffusion coefficient can be calculated from:

$$D \sim \frac{r_0^2}{4\tau_D}$$

Equation 2-39

As only a few molecules are in the focal volume at any time, the signal-to-background ratio can be quite low. When calculating the count rate of a molecule this must be considered.[26] This is especially crucial for molecules with low brightness, such as fluorescent base analogues. For an ideal situation, where background is equal to zero, counts per molecule (η) can be written as:

$$\eta = \frac{\langle F(t) \rangle}{N}$$

Equation 2-40

To correct for a non-correlating background $\langle F(t) \rangle$ can be scaled accordingly:

$$\frac{\langle F(t) \rangle^2}{[\langle F(t) \rangle - \langle F_{BG}(t) \rangle]^2}$$

Equation 2-41

Where $\langle F_{BG}(t) \rangle$ is background count rate.

2.4 Computational Chemistry

The computation background discussed here will focus predominantly on that relating to the calculation of molecular structure and energy. The majority of the section will discuss density functional theory, as this is the main technique utilised throughout this thesis.

The goal of a computational chemist is to describe the properties of molecules. Specifically, the energy of a molecule and how that energy changes with varying atomic position. Electrons, due to their small size and mass, cannot be described using classical Newtonian physics but instead must be considered in terms of quantum mechanics and the Schrödinger equation. Methods that endeavour to solve the Schrödinger equation are referred to as *ab initio*, meaning from first principles.

In its simplest form the Schrödinger equation takes the form:

$$\hat{H}\Psi = E\Psi$$

Equation 2-42

where $\hat{H}$ is the Hamiltonian operator, a term which characterises the total energy of a given system; this includes the kinetic energy of the electrons and nuclei as well as potential energy associated with the repulsion between nuclei, the repulsion between electrons and the attraction between the two. E is energy and Ψ is a wavefunction.

There are certain circumstances under which the Schrödinger Equation can be solved exactly. For the systems of interest to most computational chemists, systems larger than one-electron systems, this is not possible. In order to solve for these large systems, it is necessary to make approximations.

The first of these approximations is the Born-Oppenheimer approximation. This states that is possible to separate the electronic and nuclear terms due to the difference in timescale between electronic and nuclear motion. Electrons, being so much lighter than nuclei, rearrange on a timescale over which the nuclei can be regarded as stationary. This approximation has been considered in writing the form of the Schrödinger equation shown in Equation 2-43. In Equation 2-43 only those terms relating to the electronic wavefunction have been included. Solving this time-independent form of the electronic Schrödinger equation for various nuclear coordinates creates a potential energy surface.

When describing the interaction of multiple electrons and nuclei it is the time-independent form of the Schrödinger equation that is used, Equation 2-43.[27] Here E represents the ground state energy of the system.

$$\left[\hat{\mathrm{H}}\Psi = \overbrace{\frac{-\hbar^2}{2m_e}\sum_{i=1}^{N}\nabla^2}^{1} + \overbrace{\sum_{i=1}^{N}V(r_i)}^{2} + \overbrace{\sum_{i=1}^{N}\sum_{j<i}U(r_i,r_j)}^{3}\right]\Psi = E\Psi$$

Equation 2-43

The Hamiltonian operator ($\hat{H}$) describes the system's kinetic and potential energy terms. N is the number of electrons. Term 1 describes the kinetic energy of each electron. Term 2 describes the potential energy of the electron-nuclei interaction and Term 3 describes the electron-

electron interaction energy. m_e denotes electron mass, and r is the spatial coordinate of each electron N. Ψ is the electronic wavefunction and is dependent on both r and N.

Separation of electron and nuclear terms is not sufficient to make the Schrödinger equation soluble, as the existence of multiple interacting electrons gives rise to a many-body problem. Further approximations are therefore required, the most basic of these is the Hartree-Fock (HF) method. The Schrödinger equation is solvable for a one electron system. If each electron in a molecular system is treated separately the total wavefunction can be written as the sum of each electron wavefunction. This allows the wavefunction to be approximated as the product of the individual electron wave functions. This is called the Hartree product. To separate wavefunctions in this manner involves ignoring interactions between electrons. This is a severe approximation which would result in unrealistically large errors. To approximate electron correlation, how movement of one electron affects the others, HF theory instead models this by assuming that each electron moves in a uniform field produced by the others.[28]

The Hartree product also violates the antisymmetry principle, the principle that a wavefunction must be anti-symmetric in respect to the exchange in position of any two electrons. This principle is based on the Pauli exclusion principle that no two fermions can be in the same state, in this case two electrons with the same spin cannot occupy the same orbital. By combining wavefunctions in a Slater determinant this violation can be overcome. The Slater determinant of N electrons consists of a NxN matrix, interchanging any two electron leads to an interchange of rows, which results in a change of sign. An example for two electrons is displayed in Equation 2-44 where χ_1 describes the spatial coordinates and x_1 describes the spin coordinates.

$$\Psi(x_1, x_2) = \frac{1}{\sqrt{2}} det \begin{bmatrix} \chi_1(x_1) & \chi_2(x_1) \\ \chi_1(x_2) & \chi_2(x_2) \end{bmatrix}$$

Equation 2-44

Hartree-Fock theory accounts for electron exchange but its approximation of electron correlation is very crude. Higher levels of theory which more accurately model correlation exist. These come with the disadvantage of increased computational cost as accuracy increases. Some of the most common approaches are configuration interaction (CI), couple cluster (CC) and Møller-Plesset (MP) perturbation theory. The MP2 method has been used for some of the work in this thesis. MP theory accounts for electron correlation by including a perturbation to the Hamiltonian operator. It is assumed that this perturbation is small compared to $\widehat{H}$. The "2" indicates this is a second order perturbation.

An alternative approach to solving the Schrödinger equation is density functional theory (DFT). This is the chief computational method utilised in the work of this thesis. The basis of DFT is two theorems derived by Kohn and Hohenberg in the mid-1960s. The result of these theorems is that a functional for energy (E[n]) can be determined from just electron density.[28] The term functional describes calculating a value from a function. Here the functional calculates energy functions from electron density, which itself is a function dependent on nuclear position.[27]

$$\text{Electron density} = \rho(x, y, z)$$

Equation 2-45

$$\text{Energy } E = F[\rho(x, y, z)]$$

Equation 2-46

This is significant as it means the energy of the system can be determined by calculating the electron density as a function of just the x, y and z electron positions. The number of variables is therefore the same, independent of the number of electrons. This allows DFT to be scaled to larger systems than HF methods without too significant a computational cost.

The total energy in terms of electron density can be described as:

$$E[\rho] = T[\rho] + E_{n-e}[\rho] + E_{ee}[\rho]$$

Equation 2-47

Where $T[\rho]$ is the electron kinetic energy, $E_{n-e}[\rho]$ is electron-nuclear attraction $E_{ee}[\rho]$ is electron-electron repulsion. This last term can be further separated into electron-electron repulsion and electron exchange. Of these terms kinetic energy and electron exchange are the most complex to solve. A method to calculate the kinetic energy term was solved by Kohn and Sham. This is done by calculating the exact energy for a system of non-interacting electrons described by Kohn-Sham orbitals (ϕ_i).

$$\rho(r) = \sum_{i=0}^{N_{elec}} \left| \phi_i \right|^2$$

Equation 2-48

An initial guess is used to calculate an initial set of Kohn-Sham orbitals, from these an electron density is calculated. This density can then be used to recalculate orbitals. Energy is minimised in this iterative fashion until a minimum is reached. The Kohn-Sham equations allow kinetic energy to be calculated but a functional is still required to describe exchange-correlation.[29]

Finding the actual functional required to successfully calculate electron-exchange is the main problem with DFT. There are three main classes of functional available, those based upon a Local Density Approximation (LDA), gradient-corrected methods and hybrid methods. The LDA is the simplest approximation and describes electron density as a homogenous electron gas. LDA works well for metallic systems but is less accurate for molecular systems. Gradient-corrected methods are an improvement on the LDA as they treat electron density as non-uniform, a more realistic approach when considering molecules. Hybrid functionals are currently some of the most commonly used. They combine aspects of exact exchange calculated from Hartree-Fock theory and the rest from empirical data or *ab initio* methods. The most widely used hybrid functional is B3LYP. This is a combination of the gradient-corrected functional LYP, which accounts for correlation, and B which accounts for the exchange.

The molecular wavefunction is constructed from atomic orbitals. A basis set constructs a molecular orbital through the linear combination of these atomic orbitals. Usually a series of Gaussian functions are used. Larger basis sets increase accuracy but incur a greater computational cost. Split valence basis sets have been used in this work. These can reduce computational cost by treating core and valence electrons separately. 6-31G* denotes 6 Gaussian functions are used to model the core. 31 denotes that two sets of functions have been used to model the valence electrons, the first using three sets of functions the second set containing one.

Basis sets can be supplemented with diffuse (+) or polarisation (*) functions. Diffuse functions are larger versions of s or p orbitals and are of importance when considering loosely bound electrons, for example where polarizability or excited states are of interest. Polarisation functions are important and always included in basis sets. Polarisation is added by including higher angular momentum functions. In order to describe, for example, the combination of two s-orbitals correctly, a p-type orbital must be included. The p-orbital accounts for the distribution of electron density along the bond being different from that perpendicular. In this same manner, d-orbitals can be included to describe p-orbitals and so forth.

Time-dependent density-functional theory (TDDFT) has also been utilised throughout this thesis to investigate excited state molecular properties. TDDFT is an expansion of DFT and its methodology is somewhat analogous. Runge and Gross proved that there is a one to one mapping between a time-dependent external potential and time-dependent electron density. Just as in DFT, time dependent Kohn Sham equations are based upon a non-interacting single

electron system. For systems that undergo a small perturbation, for example most spectroscopy applications, linear-response theory is used.

2.5 DNA

2.5.1 DNA Structure

The deoxyribose nucleic acid (DNA) double helix is one of the most widely recognisable biomolecular structures. It consists of two intertwining polynucleotide strands which run antiparallel to each other. Each strand is made up of repeating nucleotide units. DNA can take multiple forms, the most common is B-form DNA, illustrated in Figure 2-9. B-form DNA, has ten base pairs per turn with 3.4 Å between neighbours in the same strand. Each full turn of the helix contains 10 bases each rotated by 36° relative to its neighbour. Base pairs refer to two complementary bases in opposite strands whereas neighbours refer to the bases in the same strand. Also illustrated in Figure 2-9 are the major and minor groove. These grooves are opposite to each other in the DNA structure and are continuous along the entire length of the strand. They exist in all DNA double helices; in B-form DNA the major groove is wider than the minor groove whereas in A-form DNA it is narrower than the minor groove but deeper.[30] These grooves provide small molecules and proteins access to the base pairs. The negatively charged phosphate groups in the backbone interact with water, whereas the minor and major groove are less polar. The final feature illustrated in Figure 2-9 is the 5' or 3' ends, these indicate directionality. The 5' end indicates the end terminating in a phosphate group attached to the 5' carbon of the sugar group and 3' indicates the OH attached to the 3' carbon of the sugar ring.

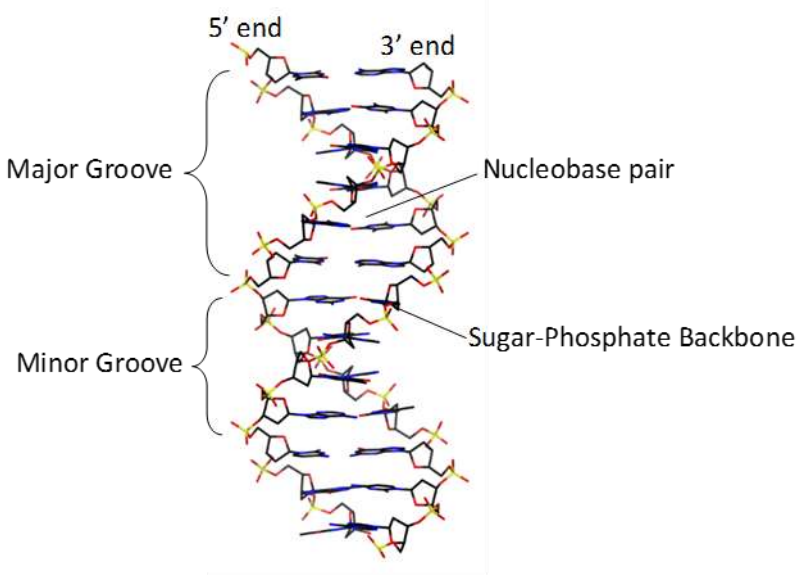


Figure 2-9. The double helix structure of B-form DNA.

It is also important to mention RNA as the molecules investigated in Chapter 4 are derivatives of uridine, a base exclusively found in RNA. RNA differs from DNA in several ways, it contains the base uracil instead of thymine and it contains a ribose sugar group in place of a deoxyribose. The difference between ribose and deoxyribose is that, as illustrated in Figure 2-10, ribose contains an additional oxygen. Although a difference of one hydroxyl group may appear to be a minor change, this results in some vast differences between DNA and RNA. The OH group makes RNA more susceptible to hydrolysis and increases the number of conformations it can adopt compared to DNA. RNA almost always exists as a single strand although, where complementary sections exist, it can fold in on itself. Transfer RNA, for example, is known to exist in a clover leaf secondary structure.

The building blocks of DNA and RNA are the bases, linked together by a sugar-phosphate backbone. The sequence of these bases is what encodes genetic information. The nomenclature used throughout this thesis when referring to DNA nucleobases is illustrated in Figure 2-10.

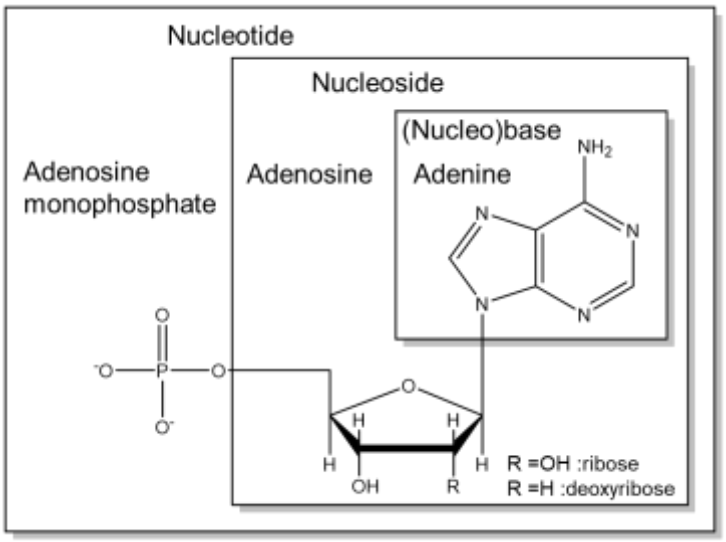


Figure 2-10. Structure and nomenclature of an adenine nucleotide, nucleoside or nucleobase. Difference between a ribose or deoxyribose sugar ring also illustrated.

The abbreviation A will be used to refer to adenine and also adenosine. Likewise T, C, G and U will be used to refer to thymine, cytosine, guanine and uracil as well as their nucleoside and nucleotide forms. These bases are illustrated in Figure 2-11. The nucleobase or free base refers to base without a sugar ring attached. The nucleoside refers to the case where it is just the base and sugar ring. Riboside or deoxyriboside refer to which sugar is attached. Nucleotide refers to where both a sugar and phosphate group are attached to the base.

The bases themselves are small aromatic molecules which absorb in the UV range: A at 257 nm with an absorption coefficient (ε_o) of $15x10^3$ $M^{-1}cm^{-1}$; C 280 nm, 13 $x10^3$ $M^{-1}cm^{-1}$; G 256

nm 12 x10^3 $M^{-1}cm^{-1}$; T 267 10 x10^3 $M^{-1}cm^{-1}$ and U 262 nm, 10 x10^3 $M^{-1}cm^{-1}$ respectively. They are virtually non-fluorescent with quantum yields in the range of ~1 x10^{-4} and picosecond or sub-picosecond lifetimes.

There are two types of base, purines and pyrimidines. The purines, adenine and guanine, base pair respectively with thymine and cytosine. In RNA, thymine is replaced by the base uracil. Hydrogen bonding between complementary bases is what holds the DNA duplex in place.

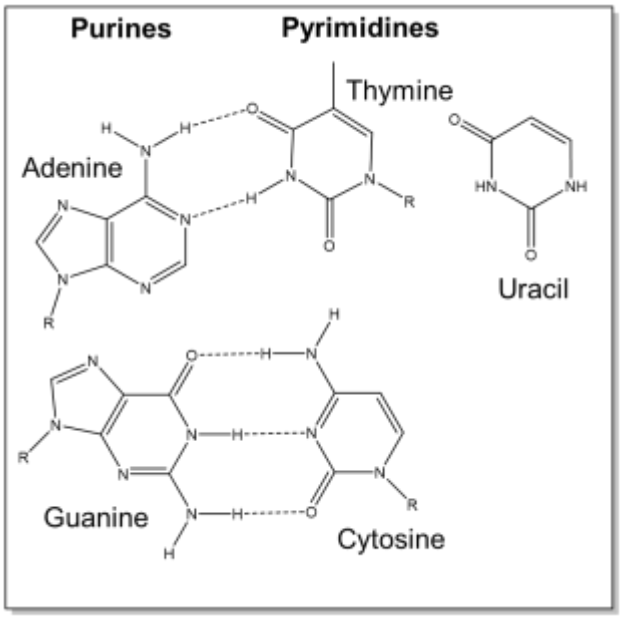


Figure 2-11. Nucleobases adenine, guanine, thymine, cytosine and uracil. A-T and G-C Watson-Crick base pairing is also illustrated.

Figure 2-11 illustrates that there are two hydrogen bonds between A and T and three between G and C. Consequently GC base pairs tend to be more stable. As well as the Watson-Crick base pairs illustrated here there are also other, less favourable base pairings that can occur. For example, a Hoogsteen base pair which involves a pair where one of the bases is rotated by 180° relative to its position in a Watson-Crick base pair. This results in a change in width of the base pair so is not compatible with B-form DNA. In a usual Watson-Crick pair both an AT and CG pair are 10.85 Å wide at their furthest point, allowing a symmetrical helix. This type of Hoogsteen base pairing has been shown to occur when DNA is bound to certain proteins or antibiotics, in AT-rich sequences or in modified or damaged oligonucleotides.[31]

Base pairing contributes to the overall stability of a double helix, although base stacking interactions between neighbours contribute more. A single incorrect base pair can decrease the stability of the helix by between ~4-16 kJ mol^{-1}.[32] If the mismatched base pair is towards the end of the strand this effect is lessened. The origin of this decrease in stability is believed to

be less favourable base stacking interactions that occur due to distortion of the geometry as a result the unmatched pair. Base stacking is a phenomenon caused by the interaction of π electron systems, which are stabilised by charge transfer and van der Waals interactions between them. In water, stacking between free bases is favourable due to the minimisation of contact of the non-polar bases with polar water. Due to the difference in the permanent dipoles of the different bases, the stacking interaction between pairs is different. The order and hence relative stability of bases can affect the flexibility of a DNA sequence. This in turn affects where, for example, a protein will interact.

When describing the structure of DNA it would be easy to think of DNA as a static structure as opposed to a dynamic biomolecule. There are various motions that occur over several different timescales. These range from motion of individual bases happening on a nanosecond timescale to slower motions such as, untwisting of the double helix or global rotations of the duplex which happen on a millisecond to second timescale.

There are two or three hydrogen bonds between complementary bases, depending on the bases as illustrated in Figure 2-11. These bonds are weak enough that even in a double helix, base-pair geometries can deviate from the arrangement shown in this Figure. For two base-paired molecules, the bases can be rotated or translated relative to each other. Bases can also undergo motion relative to neighbouring bases in the same strand and these motions are described in terms of shift, slide, rise, tilt, roll and twist, as illustrated in Figure 2-12.

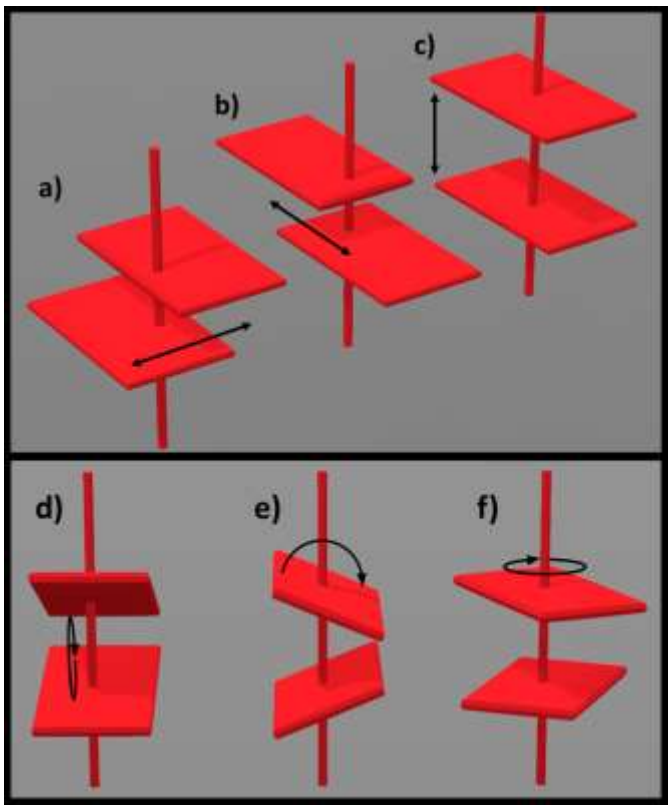


Figure 2-12. Relative motion of neighbouring bases. a) Shift b) Slide c) Rise d) Tilt e) Roll f) Twist

As well as the dynamics of the bases, there are global motions of the DNA or RNA strand. A single strand has a high degree of flexibility due to rotations of bonds in the backbone that can occur. A double helix, although more rigid, can also curve, bend, stretch or twist. These properties can be affected by environmental factors such as salt concentration.

Both this global and local flexibility is important when considering the interaction between DNA and other molecules. In order to be repaired, for example, a DNA base must be flipped out of the helix and into an enzyme. This requires a degree of flexibility. Both structure and flexibility are important factors when considering the binding of small molecules such as drugs to DNA.

2.5.2 Fluorescent Labelling of DNA

Fluorescent dyes are widely used in the study of DNA due to the low intrinsic fluorescence of DNA itself. The type of fluorophore used depends on the application, for example DNA sequencing or microscopy. The most common method of attaching a fluorescent dye to DNA is to covalently bind it, *via* a linker, to the backbone. This can be done with virtually any dye, for example rhodamines or Alexa dyes, bright dyes that can be detected at a single-molecule level. The main disadvantages of attaching a dye in this manner are that it can interact with

DNA by end stacking or intercalating and can interfere with DNA protein interactions. The other disadvantage of attaching a dye *via* a linker is that the orientation of the dye is not fixed and is unknown. Other types of commonly used fluorescent dyes include groove-binding and intercalating. These dyes are non-covalently bonded to DNA. Intercalating dyes are usually flat, aromatic molecules which fit between the DNA base pairs. Some examples include acridine orange and ethidium bromide, they have both been used extensively as nucleic acid stains in cells where they can permeate the cell membrane. Groove-binding dyes attach to either the major or minor groove, depending on preferred electrostatics. These include 4',6-diamidino-2- phenylindole (DAPI) and Hoechst dyes. The binding of these dyes causes an elongation of the helical structure which can affect DNA dynamics.

2.6 Fluorescent Base Analogues

An alternative to attaching dyes externally to DNA is to make internal modifications.[33] This involves replacing one or more natural bases with modified versions. There are several methods available to do this, most commonly solid-phase assisted synthesis. A FBA has the advantage over an external probe that it is directly inserted into the helix where it can act as a probe of structure and interaction from within the duplex, as opposed to an external dye which is further removed. A second advantage is that the position in the helix and its orientation are well defined. An ideal FBA will not cause perturbation to the helix structure and thus native function will be preserved. This is an advantage over a dye attached *via* a linker which can insert within the major or minor groove and perturb the structure.

There are several techniques that are used to investigate the effect of incorporation of a FBA into a duplex. Most commonly used are melting temperature measurements and circular dichroism. If a duplex is heated to ~ 90°C there is enough kinetic energy to break all hydrogen bonds between the strands, denaturing the DNA. The melting temperature is the temperature at which 50% of the duplexes become denatured. Double stranded DNA absorbs UV light less strongly (by ~10-20%) than single stranded DNA. By measuring absorbance as a function of temperature, the melting temperature can be determined. Melting temperature depends on the length of a duplex, the sequence and also the environment (pH etc.). A lower melting temperature indicates a less stable strand. By comparing the melting temperature of a duplex containing a FBA to the native duplex, the stabilising or destabilising effect of the FBA can be ascertained. Circular dichroism (CD) is a technique that reports on the secondary structure of biomolecules.[34] CD is the difference in absorption by a molecule of left-handed circularly polarised light and right-handed. In DNA, CD arises due to the relative orientation of bases in

the helical structure. The CD spectrum of a modified and unmodified strand can be compared, to elucidate if a FBA is causing perturbation to the global duplex structure.

There are several areas in which FBAs have found use. FBAs that are sensitive to their microenvironment can report on properties such as polarity of the major and minor grooves, and structural information such as conformational changes or hybridization. FBAs can also report on the interaction of DNA with enzymes by reporting on base flipping or depurination and depyrimidination. Other interactions, such as those between DNA and ligands, have also been investigated using FBAs.

In the context of this thesis a fluorescent base analogue (FBA) will be used to refer to a molecule that both significantly resembles a natural base in size and shape and has the ability to form hydrogen bonds with a complementary base. The properties and usage of some of the more well studied fluorescent base analogues will now be discussed.

2.6.1 2-Aminopurine

Since the discovery of its fluorescence in 1969, 2-aminopurine (2AP) has become the most widely used fluorescent base analogue.[35,36] The 2AP structure is shown in Figure 2-13.

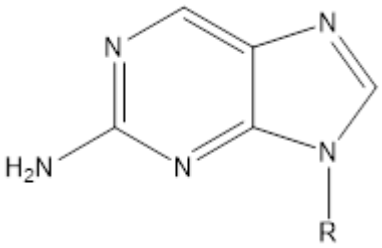


Figure 2-13. Structure of 2-aminopurine (2AP). R represents the sugar moiety.

2AP is an analogue of adenine, differing only in the position of the amine group. It can form a Watson-Crick base pair with thymine and a less stable, wobble base pair with cytosine. An absorption band at 305 nm allows 2AP to be selectively excited, away from the absorption of the natural bases. 2AP has a high quantum yield (0.68) and a long, single exponential lifetime (10.6 ns) as a free base in water. Although typically reported as a single exponential, by globally fitting the fluorescence decay of free 2AP in water, two lifetimes can be resolved and are due to the presence of 9H and 7H tautomers.[37,38] This is not true of the 2AP riboside because the presence of the ribose substituent at N9 prevents tautomerism. The quantum yield of 2AP varies depending on solvent, as does the wavelength of emission. With decreasing polarity, there is a blue shift in emission spectrum and a decrease in quantum yield.[39] This polarity-dependence is due to the relative stability of a $n\pi^*$ and $\pi\pi^*$ state. Intersystem crossing

to the $n\pi^*$ is the main pathway for non-radiative decay.[38] In polar solvent, the $n\pi^*$ state is higher in energy than in less polar solvents.

When incorporated into DNA, 2AP undergoes significant quenching with the quantum yield decreasing at least 100 fold.[1] As well as a shorter average lifetime, the fluorescence decay becomes multiexponential. This is due to the multiple conformations 2AP can adopt in DNA and will be discussed in greater detail in Chapter 6. The main advantage of different conformations having different fluorescence lifetimes is that information about population can be gained, for example the proportion of bases that are base-flipped and interacting with an enzyme. The environmental sensitivity displayed by 2AP is enormously useful for probing DNA-protein interactions, with 2AP being used extensively in studies into DNA structure and dynamics.[35] Some recent examples of the applications of 2AP include its use to study melting of a four-way junction, 2AP as part of an assay to detect the DNA glycosylases hOGG1 and UDG, and as a probe to investigate the folding of a RNA-cleavage DNAzyme.[40,41,42]

2AP has limitations and these have led to the need for the development of further base analogues. Although the quenching 2AP undergoes in DNA can be an asset, the extent to which it occurs can be limiting. In DNA the low quantum yield means that 2AP is not bright enough to be detected at a single-molecule level. Measuring a low quantum yield can be challenging and consequently quantifying changes in brightness can be hard. There are also situations in which the different 2AP stacking orientations can be disadvantageous. In an oligonucleotide, most 2AP is well-stacked and has a very short lifetime and low quantum yield, while a small percentage is extra-helical and has a higher quantum yield. This extra-helical population becomes disproportionately represented in steady-state measurements which can lead to erroneous conclusions about the system of study. This was recently shown to be the case for the primer binding site of HIV-1.[43] Another shortcoming of 2AP is its excitation and emission wavelengths. 2AP requires UV excitation and emits in the UV ~370 nm, making it unsuitable for use in microscopy. New FBAs attempt to improve upon 2AP by retaining high quantum yields in DNA and having more red-shifted absorption and emission spectra. To increase the range of use, analogues of more of the natural bases are also being developed. A more comprehensive review of 2AP photophysics and use of 2AP as a probe of DNA structure and DNA enzyme interaction has been published by Jones and Neely.[35]

2.6.2 The Pteridines

Developed by Hawkins and Pfleiderer in the late 1990's, the pteridines are a class of guanine and adenine analogues. The most successful of these were the adenine analogues 6-MAP and DMAP, and guanine analogues, 3-MI and 6-MI, all of which are now commercially available.

The structures of these FBAs are shown in Figure 2-14. 6MAP, DMAP and 3MI destabilise an oligonucleotide to about the same degree as having a single base pair mismatch. 6-MI forms a base pair with cytosine and does not destabilise a duplex. The absorption maxima of the pteridines occur at ~350 nm with emission ~440 nm. Both absorption and emission are red shifted compared to 2AP.

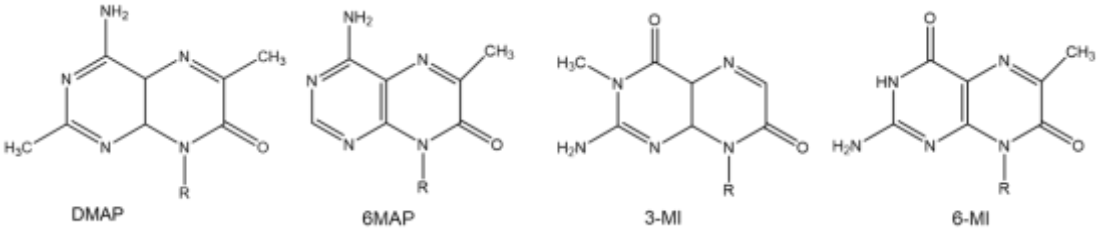


Figure 2-14. Structure of pteridines DMAP, 6MAP, 3-MI and 6-MI. R represents the sugar moiety.

As free bases, the pteridines have high quantum yields: 6-MAP one of 0.39; DMAP, 0.48; 3-MI, 0.88 and 6-MI that of 0.70. Much like 2AP, the quantum yield is significantly quenched in DNA due to base stacking interactions. The degree of quenching is sequence specific, the quantum yield of DMAP varies from <0.01 to 0.04, 3-MI from 0.03 to 0.29, 6-MI from 0.63 to 0.02. Multiexponential decay is observed for the pteridines in DNA. This will be discussed further in Chapter 6.

The intensity change between the free base and stacked has been exploited in a similar way to 2AP to investigate DNA-protein interactions. 6MAP, amongst other studies, has been used to investigate DNA base-flipping by the flavoprotein photolyase and the premelting transition of A-tracts in DNA, in various A-tract-containing sequences.[44,45] 3-MI has been used to monitor the DNA-repair enzyme O^6-Alkylguanine, and cellular uptake of oligonucleotides.[46,47] In the latter study it was shown that this was more accurate than using a fluorescein label, as uptake by the nucleic acid conducting channel was hindered by oligonucleotides labelled with fluorescein. Recently 6-MI has found use as a FRET donor. With fluorescein as a acceptor, FRET-based measurements were made in order to inform the structure of a Histone-like U93 protein-DNA complex.[48] Both 6-MI and 3-MI have been used in quadruplex priming amplification, to signal quadruplex formation.[49,50] 6MAP has found use as a probe of base-flipping by DNA photolyase.[51]

2.6.3 Fluorescent Base Analogues Developed by Tor

The group of Yitzhak Tor have made significant contributions to the field of isomorphic fluorescent base analogues including those FBAs studied in Chapter 4 of this thesis. One of

their major contributions came in 2011, a fluorescent RNA alphabet comprising, an analogue of each RNA base. These were named thA, thC, thU and thG in reference to the thiophene ring they contain[52]. These analogues are illustrated in Figure 2-15.

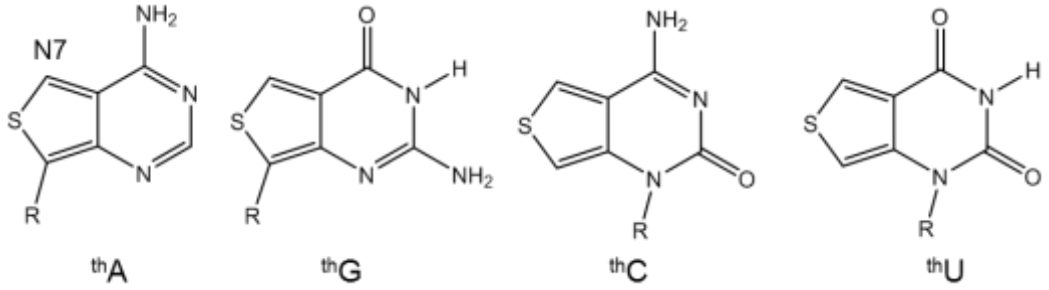


Figure 2-15. Structure of thA, thG, thC and thU. R represents the ribose sugar moiety.

These analogues were found to have absorption maxima between 304-345 nm, solvent-dependent quantum yields and emission in the visible range (378 – 453 nm). thG appeared to be the most promising candidate with absorption at 321 nm in water, 333 nm in dioxane; emission at 453 nm in water, 424 nm in dioxane and a QY of 0.46 in water and 0.50 in dioxane. It was found to retain a relatively high, but quenched, QY in an oligonucleotide (0.1) and, from melting temperature measurements, thG containing oligonucleotides were shown to have a similar stability to their unmodified counterparts. thG has been used in several studies; for example to investigate methylcytosine flipping, mediated by the set and ring associated domain of protein UHRF1.[53]

Recently, modified versions of the emissive RNA alphabet have been synthesised. These analogues have been modified by the inclusion of a N at the position labelled N7, Figure 2-15. This modification was made to improve hydrogen bonding ability of thA and thG. This new, modified family are named with the superscript tz in place of th. All analogues were found to be highly sensitive to pH and polarity.[54]

A second structural motif employed by the Tor group is that of isomorphic 5-substituted uridines shown in Figure 2-16. At the X position either a CH group or a nitrogen was inserted; at the Y position either a sulphur or oxygen.

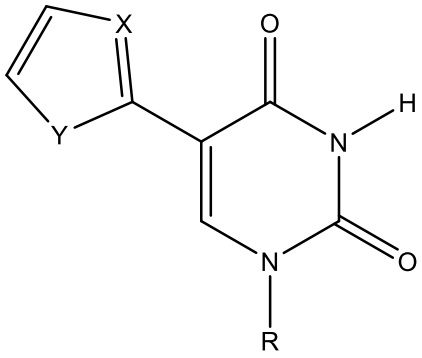


Figure 2-16. Isomorphic 5-substituded uridines differing by the substituent at X and Y. A) X= CH, Y=O; B) X=CH, Y=S; C) X= N, Y=O or D) X= N, Y=O.

The quantum yields exhibited by these molecules were found to be <0.1 to 0.03. This is higher than those of the natural bases but still very low compared to FBA such as 2AP. These compounds were shown to be sensitive to polarity, pH and viscosity[55,56]. Compound A, CH at position X and O at position Y, is the most widely used of this series. It has the highest quantum yield 0.03 in water and showed the largest sensitivity to polarity, manifesting as a blue shift in emission wavelength with decreasing polarity and a decrease in quantum yield. Its absorption maximum in water is similar to 2AP, 314 nm, and its emission is red shifted, 430 nm. These FBA act as thymine analogues with the melting temperature of an oligonucleotide containing compound A shown to be the same as that of an unmodified duplex. Amongst other things Compound A has been used to measure the polarity of the major groove and to investigate the interaction between TAR-Tat and HIV-1.[57,58] This structural motif, a 2′-deoxy-U core conjugated to an aromatic five-membered heterocycle is the same design implemented in the compounds discussed in Chapter 4.

2.6.4 Fluorescent Base Analogues Developed by Wilhelmsson

Quadracyclic adenine, qA, is a fluorescent base analogue produced by the Wilhelmsson Group is a precursor to the FBAs discussed in Chapter 5 of this thesis. Its structure is shown in Figure 2-17.

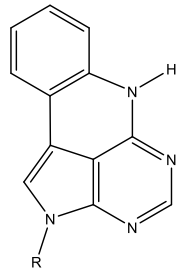


Figure 2-17. Structure of quadracyclic adenine qA.

Similar to 2AP, qA is an analogue of adenine. The absorption maximum of qA is red-shifted compared to that of 2AP, 335 nm in water. Emission occurs at 456 nm in water. In methanol the absorption maximum is slightly red-shifted compared to in qA in water and emission is blue-shifted to 340 nm. The emission spectral profile is independent of excitation wavelength but quantum yield varies. The quantum yield of qA in water is lower than that of 2AP, only 0.1 in water when excited at 377 nm and even lower when excited at shorter wavelengths, 0.06 when excited at 295 nm. Quantum yield is also excitation-wavelength-dependent in methanol, varying from 0.10 when excited at 295 nm to 0.12 when excited at 377 nm. Despite this excitation-dependence of quantum yield, time-resolved measurements made by exciting at 295 nm, 300 nm and 377 nm were reported to be single exponential with a lifetime of 3.2 ns, in all cases. An explanation for this behaviour was given as the presence of a "parasite state" between the S_2 and S_1 states. When excitation occurs to the S_2 state a portion of the excited population decays *via* a non-radiative pathway which is not accessible when excitation occurs the S_1 state.

The quantum yield of qA in oligonucleotides was found to be higher than that of 2AP in some sequences, between 0.006 and 0.002. No mechanism has been suggested for this quenching. The highest quantum yields were observed when qA was neighboured by two purines, AA, GG or GA. Lifetime measurements were also made for qA in a series of sequences and these decays were fitted to either two or three components, depending on sequence. In single strands, one lifetime component of between 3-4 ns (similar to that of free qA) accounts for between 68 to 8% of the population a second component of 0.4-0.6 ns accounts for between 8- 30% and a third component 0.0-0.1 ns accounts for the remaining 20 - 70%. The lifetimes remain fairly consistent irrespective of sequence but the proportion of each varies significantly. In a double strand the lifetimes are similar to those in a single strand but the proportion of each varies.

2.6.5 tC Family

tC, a cytosine analogue, differs from other fluorescent base analogues discussed here in that it is environmentally insensitive. tC has an absorbance maximum around 375 nm and a quantum yield of 0.20, and forms a stable base pair with guanine and also adenine. This quantum yield is relatively unchanged on inclusion in an oligonucleotide, ranging from 0.17 to 0.24 depending on the identity of nearest neighbours. The absorption and emission maxima in DNA are also largely independent of neighbouring base, with absorption consistently at ~395 nm and emission at ~510 nm in both single and double-strands. The lifetime is longer in oligonucleotides than in water, 3.8 ns in water, but 5.3 to 6.5 ns for single-strands, and 5.7 to 6.9 ns for double-strands. The single exponential decay in DNA reported by Wilhelmsson *et al.* implies that tC adopts just one conformation in DNA, not multiple conformations as

observed for 2AP. There is, however, evidence that the decay is not a single exponential but there are in fact two components, a ~3 ns component that accounts for ~20% of the population and a longer lifetime component, 5-6 ns.[59] Tautomerisation at the hydrogen bonding face was suggested as a possible source of the second species, as this also explains the ability of tC to base pair with both guanine and adenine.

A series of environmentally insensitive tC derivatives with differing absorption and emission properties have been synthesised and are shown in Figure 2-18.

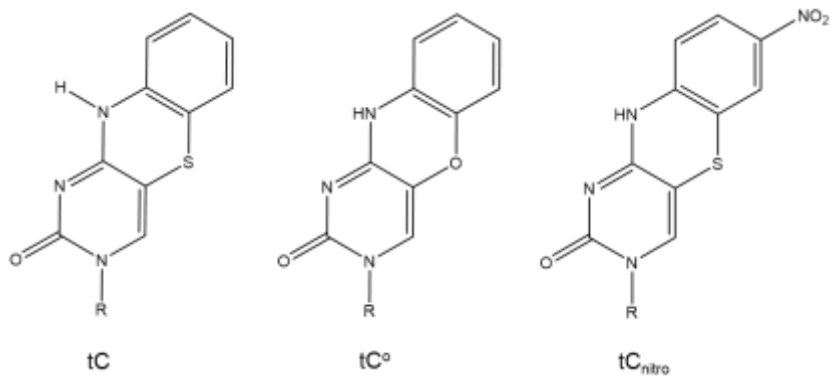


Figure 2-18 Structure of tricyclic cytosine, tC, tC° and tC$_{nitro}$. R indicates sugar.

The unresponsiveness of tC makes it ideal for use in FRET measurements. Insensitive base analogues are important for FRET, as changes in intensity can be attributed to changes in distance between donor and acceptor, not due to environmental change. FBA FRET pairs have the advantage over labels attached by a linker as their position is well determined. A tC^{O} donor, tC$_{nitro}$ acceptor (Figure 2-18) FRET pair has been successfully used to study a salt-induced B-DNA to Z-DNA transition.[60]

tC is slightly larger than the natural bases and the other fluorescent base analogues discussed so far. For this reason is it sometimes referred to as a size-expanded base. NMR, melting temperatures and circular dichroism (CD) measurements on modified and non-modified duplexes indicated that, tC, tC° and tC$_{nitro}$ do not perturb the duplex structure despite this larger size.[61,62,63]

The FBAs mentioned here are some of the most commonly used, commercially available FBAs and those that will be relevant to the molecules studied throughout this thesis. It has been shown that a wide range of FBA properties are useful and important in the study of DNA. Environmental sensitivity where emission is quenched in a duplex and less quenched when base flipped can be useful as a probe of base stacking. However, in other situations

environmental insensitivity can also be useful, for example in FRET applications. The FBAs discussed here do not constitute all in existence, there are hundreds that have been produced and there are several reviews that contain more expansive coverage. A review by Tor *et al.* contains an extensive list (100+) of the most important emissive analogues of nucleosides and nucleotides and an overview of their use.[64] A review by Wilhelmsson covers the photophysical properties of the most widely used FBAs with a specific focus on environmentally insensitive FBAs.[65] A recent review by Kool *et al.* looks at the history of canonical (retains ability to form a Watson-Crick base pair) and non-canonical FBA development, as well as labelling techniques and synthetic strategy.[66]

From the FBAs reviewed here, areas of future improvement can be elucidated. The main aims of future FBA development can be divided into three parts. Firstly, red–shifting absorption and emission further into the visible, thus opening up the prospect of microscopy. Secondly, increasing quantum yields of environmentally sensitive probes in order that measurements can be more easily and widely implemented. Both these tasks must be done while retaining the isomorphic structures that allow insertion into DNA without duplex perturbation. It would also be advantageous to increase brightness in order to open up the possibility of single-molecule techniques. Finally, an increased understanding of fluorescence properties, including quenching mechanisms, is needed in order to inform future FBA design.

2.7 References

1 J. R. Lakowicz, *Principles of Fluorescence Spectroscopy*, Springer, 3rd edn., 2006.

2 P. Atkins and J. de Paula, *Atkins' Physical Chemistry*, Oxford University Press, Oxford, 8th edn., 2006.

3 B. R. Masters and P. T. C. So, *Handbook of Biomedical Nonlinear Optical Microscopy*, Oxford University Press, New York, United States, 1st edn., 2008.

4 R. K. Neely, PhD Thesis, The University of Edinburgh, 2005.

5 N. Boens, W. Qin, N. Basaric, J. Hofkens, M. Ameloot, J. Pouget, J. Lefevre, B. Valeur, E. Gratton, N. D. Silva, M. VandeVen, Y. Engelborghs, K. Willaert, A.

Sillen, A. J. W. G. Visser, A. Van Hoek, J. R. Lakowicz, H. Malak, I. Gryczynski, A. G. Szabo, D. T. Krajcarski, N. Tamai and A. Miura, *Anal. Chem.*, 2007, **79**, 2137–2149.

6 G. S. He, L. S. Tan, Q. Zheng and P. N. Prasad, *Chem. Rev.*, 2008, **108**, 1245–1330.

7 C. Xu and W. W. Webb, *J. Opt. Soc. Am. B*, 1996, **13**, 481.

8 M. A. Albota, C. Xu and W. W. Webb, *Appl. Opt.*, 1998, **37**, 7352.

9 M. Drobizhev, S. Tillo, N. Makarov, A. Rebane and T. E. Hughes, *J. Phys. Chem. B Lett.*, 2009, **113**, 855–859.

10 M. Drobizhev, N. S. Makarov, S. E. Tillo, T. E. Hughes and A. Rebane, *Nat. Methods*, 2011, **8**, 393–399.

11 H. Y. Woo, B. Liu, B. Kohler, D. Korystov, A. Mikhailovsky and G. C. Bazan, *J. Am. Chem. Soc.*, 2005, **127**, 14721–14729.

12 M. Johnsen and P. R. Ogilby, *J. Phys. Chem. A*, 2008, **112**, 7831–7839.

13 M. Wielgus, R. Zaleśny, N. A. Murugan, J. Kongsted, H. Ågren, M. Samoc and W. Bartkowiak, *ChemPhysChem*, 2013, **14**, 3731–3739.

14 M. Wielgus, W. Bartkowiak and M. Samoc, *Chem. Phys. Lett.*, 2012, **554**, 113–116.

15 N. Arul Murugan, J. Kongsted, Z. Rinkevicius, K. Aidas, K. V. Mikkelsen and H. Ågren, *Phys. Chem. Chem. Phys.*, 2011, **13**, 12506.

16 M. Pawlicki, H. A. Collins, R. G. Denning and H. L. Anderson, *Angew. Chemie Int. Ed.*, 2009, **48**, 3244–3266.

17 C. Toepfer and J. R. Sellers, in *Fluorescent Methods for Molecular Motors*, eds. C. P. Toseland and N. Fili, Springer, 2014, vol. 105, pp. 193–210.

18 T. Hirschfeld, *Applied Optics*, 1976, **15**, 2965–2966.

19 E. Shera, N. K. Seitzinger, L. M. Davis, R. A. Keller and S. A. Soper, *Chem. Phys. Lett.*, 1990, **174**, 553–557.

20 T. Schmidt, G. J. Schütz, W. Baumgartner, H. J. Gruber and H. Schindler, *Proc. Natl. Acad. Sci.*, 1996, **93**, 2926–2929.

21 T. Ha and P. Tinnefeld, *Annu. Rev. Phys. Chem.*, 2012, **63**, 595–617.

22 M. Honda, J. Park, R. A. Pugh, T. Ha and M. Spies, *Mol. Cell*, 2009, **35**, 694–703.

23 A. Gräslund, R. Rigler and J. Widengren, *Single molecule spectroscopy in chemistry, physics and biology*, Springer, 2010.

24 C. Flors, C. N. J. Ravarani and D. T. F. Dryden, *ChemPhysChem*, 2009, **10**, 2201–2204.

25 F. Persson, P. Bingen, T. Staudt, J. Engelhardt, J. O. Tegenfeldt and S. W. Hell, *Angew. Chemie - Int. Ed.*, 2011, **50**, 5581–5583.

26 S. T. Hess, S. Huang, A. A. Heikal and W. W. Webb, *Biochemistry*, 2002, **41**, 697–705.

27 D. S. Sholl and J. A. Steckel, in *Density Functional Theory: A Practical Introduction.*, John Wiley & Sons, Inc, 2009, pp. 1–33.

28 C. Morrison, *Chemistry 2 Lecture Notes*, 2011.

29 N. M. Harrison, *Am. J. Phys.*, 1995, **2**, 1–26.

30 E. Kendrew, J, Lawrence, *Encyclopedia of molecular biology*, Oxford ; Cambridge, Mass., USA: Blackwell Science., 1994.

31 E. N. Nikolova, H. Zhou, F. L. Gottardo, H. S. Alvey, I. J.Kimsey and H. M. Al-Hashimi, *Biopolymers*, 2013, **99**, 955–968.

32 C. Thomas E, in *Biophysical Chemistry of Nucleic Acids and Proteins*, Helvetian Press, 2010, pp. 49–55.

33 S. Matsika, in *Photoinduced Phenomena in Nucleic Acids I: Nucleobases in the Gas Phase and in Solvents*, eds. M. Barbatti, A. C. Borin and S. Ullrich, Springer International Publishing, Cham, 2015, pp. 209–243.

34 G. R. Bishop and J. B. Chaires, in *Current Protocols in Nucleic Acid Chemistry*, John Wiley & Sons, Inc., 2001.

35 A. C. Jones and R. K. Neely, *Q. Rev. Biophys.*, 2015, **2**, 1–36.

36 K. B. Hall, *Methods Enzymol.*, 2009, **469**, 269–85.

37 R. K. Neely, S. W. Magennis, D. T. F. Dryden and A. C. Jones, *J. Phys. Chem. B*, 2004, **108**, 17606–17610.

38 R.-X. He, X.-H. Duan and X.-Y. Li, *Phys. Chem. Chem. Phys.*, 2006, **8**, 587–591.

39 K. Evans, D. Xu, Y. Kim and T. M. Nordlund, *J. Fluoresc.*, 1992, **2**, 209–216.

40 C. E. Carr and L. A. Marky, *J. Am. Chem. Soc.*, 2017, **139**, 14443–14455.

41 Y. Zhang, C. C. Li, B. Tang and C. Y. Zhang, *Anal. Chem.*, 2017, **89**, 7684–7692.

42 R. Saran, L. Yao, P. Hoang and J. Liu, *Biochimie*, 2017, 1–6.

43 M. Sholokh, R. Sharma, D. Shin, R. Das, O. A. Zaporozhets, Y. Tor and Y. Mély, *J. Am. Chem. Soc.*, 2015, **137**, 3185–3188.

44 K. Yang, S. Matsika and R. J. Stanley, *J. Phys. Chem. B*, 2007, **111**, 10615–10625.

45 K. E. Augustyn, K. Wojtuszewski, M. E. Hawkins, J. R. Knutson and I. Mukerji, *Biochemistry*, 2006, **45**, 5039–5047.

46 A. M. Moser, M. Patel, H. Yoo, F. M. Balis and M. E. Hawkins, *Anal. Biochem.*, 2000, **281**, 216–222.

47 J. A. Costa, E. Leal-Pinto, S. C. Henderson, T. Zabel, M. E. Hawkins and B. Hanss, Pteridines, 2014, **23**, 81–89.

48 A. T. Moreno, PhD thesis, Wesleyan University, 2014.

49 J. Johnson, R. Okyere, A. Joseph, K. Musier-Forsyth and B. Kankia, *Nucleic Acids Res.*, 2013, **41**, 220–228.

50 S. Gogichaishvili, L. Lomidze and B. Kankia, *Anal. Biochem.*, 2014, **466**, 44–48.

51 K. Yang and R. J. Stanley, *Photochem. Photobiol.*, 2008, **84**, 741–749.

52 D. Shin, R. W. Sinkeldam and Y. Tor, *J. Am. Chem. Soc.*, 2011, **133**, 14912–14915.

53 V. Kilin, K. Gavvala, N. P. F. Barthes, B. Y. Michel, D. Shin, C. Boudier, O. Mauffret, V. Yashchuk, M. Mousli, M. Ruff, F. Granger, S. Eiler, C. Bronner, Y. Tor, A. Burger and Y. Mély, *J. Am. Chem. Soc.*, 2017, **139**, 2520–2528.

54 A. R. Rovira, A. Fin and Y. Tor, *Chem. Sci.*, 2017, **8**, 2983–2993.

55 N. J. Greco and Y. Tor, *J. Am. Chem. Soc.*, 2005, **127**, 10784–10785.

56 R. W. Sinkeldam, A. J. Wheat, H. Boyaci and Y. Tor, *ChemPhysChem*, 2011, **12**, 567–570.

57 R. W. Sinkeldam, N. J. Greco and Y. Tor, *ChemBioChem*, 2008, **9**, 706–709.

58 S. G. Srivatsan and Y. Tor, *Tetrahedron*, 2007, **63**, 3601–3607.

59 R. S. K. Lane and S. W. Magennis, *RSC Adv.*, 2012, **2**, 11397.

60 B. Dumat, A. F. Larsen and L. M. Wilhelmsson, *Nucleic Acids Res.*, 2016, **44**.

61 K. Börjesson, S. Preus, A. H. El-Sagheer, T. Brown, B. Albinsson and L. M. Wilhelmsson, *J. Am. Chem. Soc.*, 2009, **131**, 4288–93.

62 K. C. Engman, P. Sandin, S. Osborne, T. Brown, M. Billeter, P. Lincoln, B. Nordén, B. Albinsson and L. M. Wilhelmsson, *Nucleic Acids Res.*, 2004, **32**, 5087–95.

63 P. Sandin, K. Börjesson, H. Li, J. Mårtensson, T. Brown, L. M. Wilhelmsson and B. Albinsson, *Nucleic Acids Res.*, 2008, **36**, 157–67.

64 R. W. Sinkeldam, N. J. Greco and Y. Tor, *Chem. Rev.*, 2010, **110**, 2579–619.

65 L. M. Wilhelmsson, *Q. Rev. Biophys.*, 2010, **43**, 159–83.

66 W. Xu, K. M. Chan and E. T. Kool, *Nat. Chem.*, 2017, **9**, 1043–1055.

Chapter 3 : Experimental

The two main experimental techniques that have been used in this thesis, steady state fluorescence spectroscopy and time-resolved fluorescence spectroscopy will be described here. Details of experimental conditions etc. will be included in specific chapters, where pertinent. Computational calculations were carried out throughout the course of this work and will also be described here. Example input and output files are included in Appendix II.

A two-photon excitation system was built specifically for the experiments carried out in this thesis and so a validation of its accuracy will be included. This will be described in Chapter 7. The single molecule techniques employed in Chapter 8 will be described in the experimental section of that Chapter.

3.1 Absorption Spectroscopy

Absorption spectra were recorded using a Cary 300 UV-visible Spectrophotometer. Optically matched, 700 μl, 10 mm path length fused silica cuvettes (Starna Scientific Ltd) were used for all measurements.

3.2 Steady-state Fluorescence Measurements

Fluorescence spectra were recorded using a Jobin Yvon Horiba FluoroMax-3 Spectrofluorimeter running FluorEssence^{TM} software v.3.5. Samples were contained in a 1 cm path length fused silica fluorescence cuvette. Concentrations were typically ~ 1 x 10^{-5} M with absorbance below 0.08 to avoid inner filter effects. Measurements were made at room temperature.

The excitation source in the Fluoromax 3P is an ozone-free xenon arc lamp with a spectral range from 240 to 700 nm. This excitation light is directed towards a monochromator and the selected wavelength is focused through the sample solution. The emitted fluorescence is collected at 90 degrees to excitation. The emitted light passes through an emission monochromator and the dispersed emission is detected by a photon-counting photomultiplier tube. The Fluoromax 3P uses Czerny-Turner monochromators, these contain a 1200 grooves/ mm grating. Typically excitation and emission slit band widths of 3 nm were used. To correct for the wavelength dependence of the lamp intensity, a small portion of the excitation light is monitored by a reference photodiode. The spectra in this thesis have been corrected accordingly.

Single point quantum yields (ϕ) were determined by comparison of the integrated fluorescence intensity to that of quinine sulphate in 0.5 M H_2SO_4 (ϕ = 0.546), unless otherwise specified, using the equation:

$$\phi_s = \phi_{Ref} \frac{I_s}{I_{Ref}} \frac{A_{Ref}}{A_s} \frac{n_s^2}{n_{Ref}^2}$$

Equation 3-1

where *I* is integrated fluorescence intensity, *A* is absorbance at the excitation wavelength, *n* is refractive index of the solvent, *ref* denotes reference and *s* sample.

3.3 Time-resolved Fluorescence Measurements

Time-resolved fluorescence spectroscopy was performed using the technique of time-correlated single photon counting (TCSPC). Measurements were made in the Collaborative Optical Spectroscopy, Micromanipulation and Imaging Centre (COSMIC) facility at The University of Edinburgh. Measurements were made at room temperature. A schematic of the system used is shown in Figure 3-1. The excitation source was a tuneable mode-locked Ti:sapphire laser (Coherent Mira 900 F) pumped by a Neodymium:Vanadate (Nd:YVO4) continuous wave solid state laser (Coherent Verdi V 10, 8 W). Pulses of duration ~ 150 fs at a repetition rate of 76 MHz were produced. A pulse picker (Coherent 9200) was used to reduce the repetition rate to 4.75 MHz: one pulse every 210 ns. The excitation light was then frequency doubled or tripled using a harmonic generator (Coherent 5-050). A fraction of each pulse was diverted to a photodiode that converts the pulse to an electronic trigger pulse, indicating the start of the measurement. When using the second harmonic this diverted light is residual fundamental and in the case of the third harmonic residual second harmonic is used.

The component of the pulse not diverted to the photodiode, passes through a neutral density filter, to control the intensity, and a Soleil-Babinet compensator, to control the orientation of the plane of the polarised excitation pulse. The Soleil-Babinet compensator is a continuously variable wave plate consisting of two quartz wedges. The angle of these wedges, and thus the thickness of the quartz that the incident beam passes through, is controlled by Edinburgh Instruments F900 software. The resolution of polariser is set such that excitation light is vertically polarised. The now vertically polarised pulse is focused into a sample, contained in a cuvette, within a sample chamber. Emitted fluorescence is collected and 90 degrees to excitation direction through a polariser set at the magic angle (54.7^{o}) relative to the vertically polarised excitation light. Use of the magic angle eliminates anisotropy effects which can bias the recorded data. A long pass filter (Schott Glass, Newport) was used to block scattered excitation light. Fluorescence emission was collected by a lens and passed through a monochromator. Slits at the entrance and exit of the monochromator were set to allow a 5 nm

bandpass. Fluorescence emission was detected by a microchannel plate photomultiplier (Hamamatsu R3809-50 series).

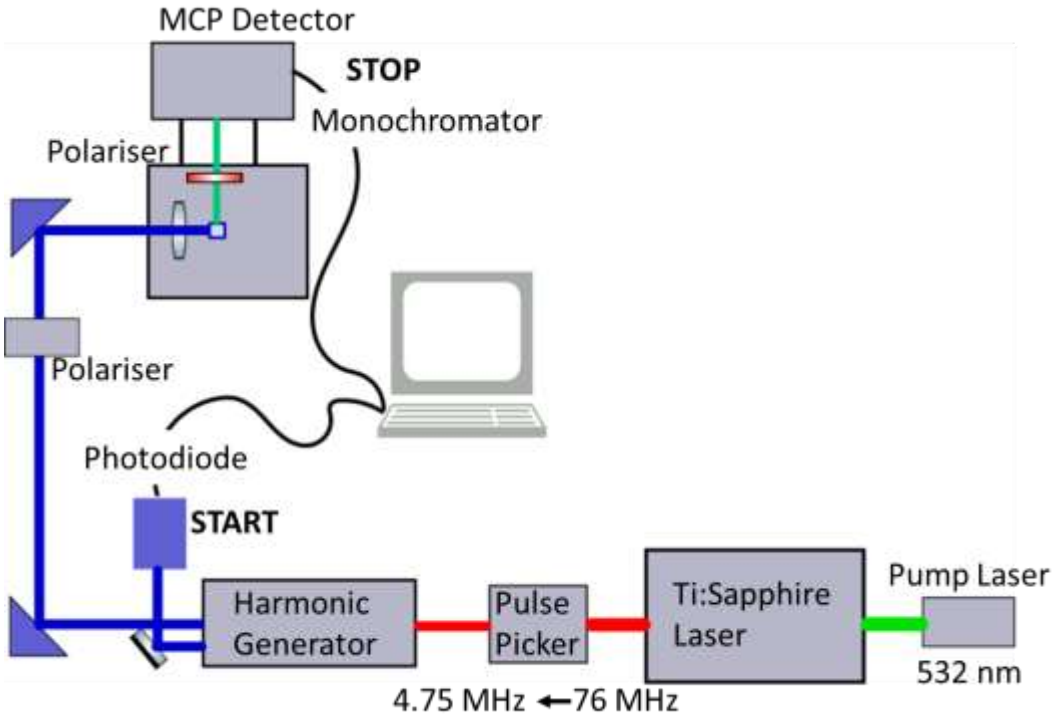


Figure 3-1. Schematic of the TCSPC set-up used for measurements of fluorescent lifetime.

Edinburgh Instruments F900 software was used to control the experimental parameters and record the fluorescence decays. Decay curves were measured over a timescale of 50 ns, across 4096 channels resulting in a resolution of 12 ps per channel. Decay curves were collected to either 10,000 or 20,000 counts in the maximum channel. Decays were typically collected at three emission wavelengths, the maximum wavelength and two wavelengths 20 nm either side of the maximum. Instrument response functions were collected, using excitation light scattered from a dilute Ludox solution, prior to each measurement. Identical conditions to those of the sample were used but with a collection count rate of ~ one tenth. The FAST software package (Edinburgh Instruments) was used for global fits of the data. Errors were taken from reconvolution fits using F900.

3.4 Computational Calculations

All calculations in this thesis were performed in Gaussian 09 which was running on the University of Edinburgh cluster Eddie.[2]

Where possible, initial starting structures were obtained from crystal structures. Where crystal structures did not exist, structures were built in ArgusLab and then optimised at a low level of

theory. Harmonic vibrational frequency calculations were performed on these optimised structures using the same level of theory and from these calculations Gibbs free energies were obtained and the nature of stationary points could be verified. Solvent effects were included using the polarisable continuum model. The input and relevant output for calculation of fluorescence emission from the first excited state are presented in Appendix II. Visualisation of molecules was carried out using Jmol.[3,4] Specific details of functionals and basis sets will be discussed where pertinent.

3.5 References

1 N. Boens, W. Qin, N. Basaric, J. Hofkens, M. Ameloot, J. Pouget, J. Lefevre, B. Valeur, E. Gratton, N. D. Silva, M. VandeVen, Y. Engelborghs, K. Willaert, A. Sillen, A. J. W. G. Visser, A. Van Hoek, J. R. Lakowicz, H. Malak, I. Gryczynski, A. G. Szabo, D. T. Krajcarski, N. Tamai and A. Miura, *Anal. Chem.*, 2007, **79**, 2137–2149.

2 Gaussian 09, Revision E.01, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox, Gaussian, Inc., Wallingford CT, 2013.

3 ArgusLab 4.0.1 M. Thompson and Planaria Software, 1997- 2004.

4 Jmol: an open-source Java viewer for chemical structures in 3D. http://www.jmol.org

Chapter 4: Photophysical Characterisation of Thiophene-6-azauridine and the Extended-6-aza-Uridines

4.1 Introduction

This chapter reports the photophysics of seven molecules: the six extended-6-azauridines, 1a-1f, and their parent compound, thiophene-6-azauridine. The structures of these molecules, as well as uridine, 6-azauridine and furan-6-azauridine, are illustrated in Figure 4-1.

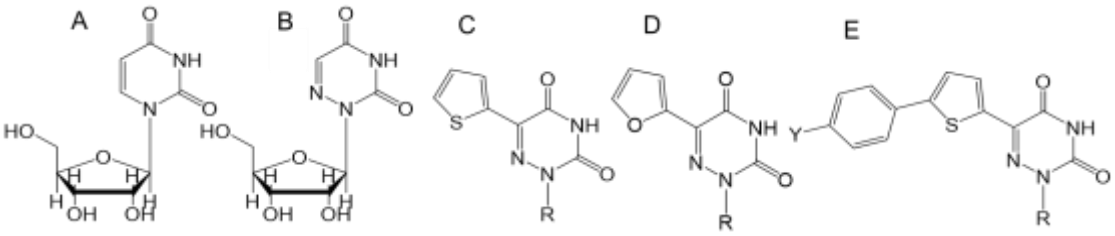


Figure 4-1. Structures of: A) uridine. B) 6-azauridine. C) thiophene-6-azauridine / parent, R=ribose. D) furan-6-azauridine R=ribose E) Extended thiophene-6-aza-uridines: 1a Y=F, 1b Y=H, 1c Y=Me, 1d Y=OMe, 1e Y=OH, 1f Y= NMe_2. R=ribose.

Produced by the Tor group, these fluorescent base analogues are the latest developments in a series of isomorphic nucleosides, based upon the addition of a 5-membered ring to uridine at the 5-position.[1] Modifications at this position on a uracil have been shown, using absorption spectroscopy and melting point measurements, to cause minimal perturbation to a duplex structure.[2] The thiophene-6-azauridine molecule discussed in this chapter differs from previous generations (described in Chapter 2 section 2.6.3) by the insertion of a nitrogen at the 6-postion of the uridine core, thus making it a 6-azauridine derivative. In addition to this modification the extended-6-azauridines contain an additional phenyl ring, with a para-substituent Y.

The extended-6-azauridines have some of the most red-shifted absorption and emission spectra of any fluorescent base analogues.[3] Emission ranges from 450 to 600 nm depending on the solvent and substituent. A popular motif for designing fluorophores with tuneable fluorescent properties is the Donor-π-Acceptor motif, a design that promotes intramolecular charge transfer. Here Tor *et al.* have managed to utilise this strategy on small molecules, whilst preserving Watson-Crick hydrogen bonding ability. The 6-azauridine moiety acts as the electron deficient acceptor, the varying substituent at position Y acts as the donor, and it is conjugated to the acceptor through the aromatic rings (Figure 4-1). The spectra of the extended 6-azauridines are red-shifted compared to the parent molecules due to this increased

conjugation. Increasing conjugation results in extended π systems, which decrease the HOMO LUMO π π* gap, resulting in longer wavelength absorption and emission. The nature of the substituent attached to this conjugated system also acts to dictate the emission wavelength. An electron donating group at the *para* position of the phenyl ring leads to an increased red-shift as a result of lowering the energy of the π* orbital. Tor *et al.* showed that there is a correlation between Hammett substituent constants, (constants based upon the difference between acid dissociation constant of benzoic acid and that of substituted benzoic acid) and absorption wavelength for these extended-6-azauridines.[3] The key photophysical properties reported previously by Tor *et al*, absorption maxima, emission maxima and quantum yield are summarised in Table 4-1.

	Solvent	Absorbance Maximum / nm	Emission Maximum /nm	Quantum Yield
Parent	Water	332	445	0.20
	Dioxane	335	415	0.80
1a Y=F	Water	357	478	0.24
	Dioxane	368	462	0.32
1b Y=H	Water	356	478	0.21
	Dioxane	371	461	0.30
1c Y=Me	Water	360	490	0.12
	Dioxane	375	468	0.38
1d Y=OMe	Water	366	525	0.02
	Dioxane	383	486	0.74
1e Y=OH	Water	381	517	<0.01
	Dioxane	385	494	0.71
1f Y=NMe_2	Water	398	484	<0.01
	Dioxane	420	575	0.20

Table 4-1. Absorption maxima, emission maxima and quantum yields of the extended 6-aza-uridines in water and dioxane. Table adapted from reference 3.

The extended-6-azauridine which has the most electron withdrawing substituent, 1a (Y=F), exhibits the most blue-shifted spectra, with an absorption maximum of 357 nm in water and an emission maximum of 478 nm. The compound with the most electron donating group, 1f (Y=NMe_2), exhibits the largest red shift with absorption at 398 nm and emission at 484 nm in water. All six nucleosides display large Stokes shifts (~ 7000 cm^{-1}) and, aside from 1f, all have greater Stokes shift in water than in dioxane. Large Stokes shift can indicate a significant difference between ground and excited state dipole moments. In the ground state, solvent molecules will be orientated in such a way as to minimise energy. On excitation this ground state solvent arrangement will be preserved. As small solvent molecules reorient rapidly, it is possible for solvent molecules to reorient around the excited-state dipole before emission occurs. This lowers the excited state energy and results in emission at longer wavelengths. Depending on the polarity of the excited state and the polarity of the solvent this solvent

relaxation will occur to different extents. For the extended 6-azauridines, water results in a greater stabilisation of the excited state than less polar dioxane, this would suggest the excited state is more polar than the ground state. A large Stokes shift is common in molecules in which excited state intramolecular charge transfer occurs, as the ground and excited state dipole moments differ significantly.

It was also found that quantum yields for all six extended-6-azauridines were higher in dioxane than water. For 1a, 1b and 1c, the quantum yield difference between water and dioxane is small. For example, 1a has a quantum yield of 0.32 in dioxane and 0.24 in water. For 1d, 1e and 1f, the quantum yield in water is significantly lower than that in dioxane. For 1d, for example, the quantum yield drops by a factor of 10 from 0.74 to 0.02. The low quantum yield of 1d, 1e and 1f in water is believed to be due to non-radiative decay pathways facilitated by hydrogen-bonding to solvent.[3]

As well as the extended-6-azauridines, the parent, thiophene-6-azauridine, will also be investigated in this Chapter. The paper that reported the synthesis of the thiophene-6-azauridine, also reported a furan derivative.[4] The steady state properties of both compounds are displayed in Table 4-2.

	Solvent	Absorbance Maximum / nm	Emission Maximum /nm	Quantum Yield
Thiophene-6-azauridine	Water	332	455	0.20
	Dioxane	335	415	0.80
	Methanol	334	433	0.50
Furan-6-azauridine	Water	320	443	0.05
	Dioxane	327	414	0.60
	Methanol	326	442	0.04

Table 4-2. Absorption maxima, emission maxima and quantum yields of the thiophene-6-azauridine and furan-6-azauridine in water and dioxane. Table adapted from reference 2.

As previously mentioned, the absorption and emission maxima of the parent thiophene-6-azauridine nucleoside are blue-shifted compared to the extended compounds. The parent absorbs at 332 nm in water, whereas 1a, for example, absorbs at 357 nm. Parent emission occurs at 455 nm in water with 1a emitting at 478 nm. Compared to the furan-6-azauradine the thiophene-6-azauridine absorption is slightly red-shifted, by ~ 10 nm in each solvent. Similar to the extended compounds, Stokes shifts for both the thiophene and furan derivatives are large (~7000 cm^{-1}), indicating excited-state charge transfer. In water and methanol, the Stokes shift is larger for the furan-6-azauradine than for the thiophene. Conversely, in dioxane it is larger for the thiophene-6-azuridine than for furan-6-azauradine. The quantum yield of thiophene-6-azauridine is higher than that of the furan derivative in all solvents. The largest difference is seen in methanol where thiophene-6-azauridine has a quantum yield of 0.50,

whereas furan-6-azauridine has a quantum yield of 0.04. Both the thiophene and furan fluorescent base analogues were shown to be sensitive to polarity and pH. When plotting the Stokes shift of thiophene-6-azauridine, in mixtures of water and dioxane, against the empirical solvent polarity parameter ($E_T(30)$/ kcal mol^{-1}) the resulting plot yielded a straight line with a slope of 0.27. As with the extended-6-azauridines, this suggests a more polar excited state than ground state. Dioxane and water were used as solvents in this study, as dioxane has a polarity similar to that of the internal environment of DNA, whereas the polarity of water is more similar to that external to the DNA duplex. Sensitivity to polarity across this range is therefore indicative of the polarity range a nucleobase will experience and can indicate if a base analogue will be a useful as a probe of, for example, DNA base stacking. The furan derivative exhibited a similar trend of increasing Stokes shift with polarity, but the sensitivity was lower.

Thiophene-6-azauridine emission was also studied as a function of pH over the range 0.05 – 12. A change in pH was found to influence both the emission maximum and the intensity of emission, but to have little effect on the absorption spectrum. Under basic conditions the thiophene-6-azauridine has a higher quantum yield and a more blue-shifted emission maximum than found under acidic conditions. A Boltzmann fit of a plot of emission intensity as a function of pH revealed the pK_a of deprotonation of 6.8, the same as that of unmodified 6-azauridine.[5] Crucially, the change was significant enough over the physiological range of pH that it was suggested that thiophene-6-azauridine would be a useful probe of RNA deprotonation.

Sensitivity of thiophene-6-azauridine and furan-6-azauridine to viscosity was also investigated in a mixed methanol/glycerol solvent. Viscosity was controlled by adjusting the ratio of methanol to glycerol. The quantum yield of the furan derivative was found to increase significantly with increasing viscosity. No such viscosity-dependence was observed for the thiophene-6-azauridine compound. It is unusual that the thiophene-6-azauridine did not show any apparent sensitivity to viscosity, as it had been previously suggested that all FBAs based on this conjugated ring structure act as molecular rotors and would therefore exhibit viscosity-dependence.[6] Molecular rotors are molecules that undergo a twisting motion in the excited state.[7] The Jablonski diagram for a typical molecular rotor such as 4,4-dimethylaminobenzonitrile (DMABN) is illustrated in Figure 4-2.[8]

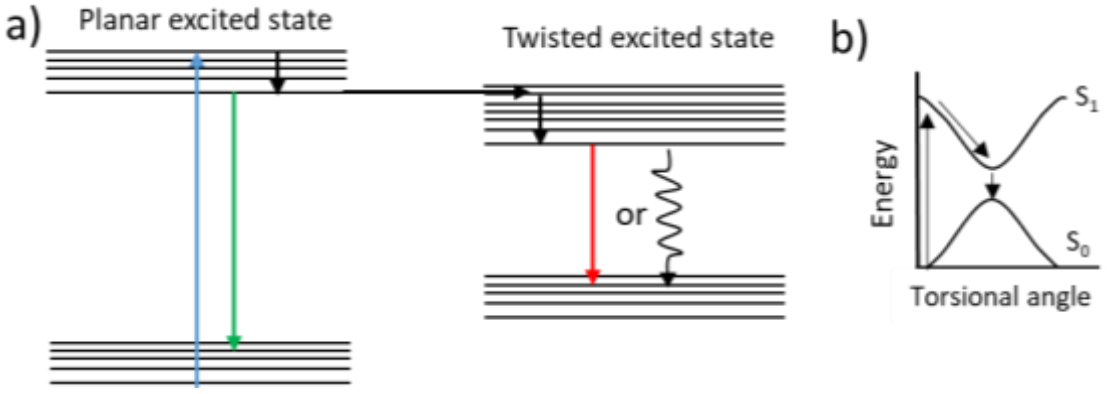


Figure 4-2. a) Jablonski diagram for a molecular rotor. Excitation occurs to the local planar excited state. Emission can occur from this local excited state or relaxation to a lower energy twisted state can occur followed by emission from this twisted state. b) Schematic of ground and excited state potential energy as a function of torsional angle.

Typically, a planar conformation of a molecular rotor is lowest in energy in the ground state. However in the excited state, a twisted conformation is most favourable.[9] Just as with any fluorophore, molecular rotors are excited from S_0 to S_1 by absorption of a photon. Vibrational relaxation to the lowest energy excited state then occurs. The difference with a molecular rotor is that the lowest energy conformation in the excited state does not have the same planar conformation as in the ground state. Instead a twisted conformation is typically lower energy. Often emission occurs from both the locally excited planar state and from the twisted state. In the ground state, the twisted conformation has a higher energy, consequently the S_1-S_0 energy gap for this twisted state is lower in energy and so its emission occurs at a longer wavelength. If this energy gap is sufficiently small, rapid non-radiative decay can occur. Depending on the relative rates of radiative and non-radiative decay from the twisted excited state, either dual emission and multiple lifetimes, or single wavelength emission and a single lifetime is observed.

The emission properties of molecular rotors tend to be highly dependent on environment. The biggest effect is seen with increasing viscosity. Increasing viscosity suppresses rotation, resulting in a greater population of the locally excited species and less of the twisted species. If the twisted species decays *via* non-radiative pathways, decreasing the proportion will result in an increase in quantum yield. If the twisted species is emissive, increase in viscosity will result in a change in emission wavelength: less of the red-shifted twisted state and more of the locally excited state. There is also often a solvent polarity effect. For example, DMABN exhibits dual emission in polar media, but single emission in non-polar solvents. This is attributed to increased stabilisation of the twisted excited state by polar solvents.

Consequently, emission from both locally excited and twisted states can occur. In less polar solvent, where the twisted state is higher in energy, only emission from the locally excited state is seen.[10]

From the thiophene-6-azauridine structure displayed in Figure 4-1(C) it can be seen that, depending on the torsional angle of the thiophene ring, there will be either a sulphur- oxygen (S-O) interaction or a sulphur-nitrogen (S-N) interaction. Non-covalent interactions between electron-deficient sulphur and electron donors, such as oxygen and nitrogen, are well-documented and often observed in X-ray crystallographic data.[11] These interactions occur as a consequence of low-lying C-S σ* orbitals on sulphur which result in "σ-holes" which possess positive electrostatic potential. These non-covalent interactions are analogous to halogen bonding. The strength of these interactions can be affected by neighbouring substituents, especially electronegative groups. When in competition, hydrogen bonding can dominate over S-O or S-N interactions. Unfavourable steric effects and other types of intermolecular interaction can also affect the favourability of a sulphur interaction. The strength of S-O bonds, unlike the strength of hydrogen bonds have been shown, using a combination of NMR and DFT calculations, to be similar irrespective of solvent.[12] It is unclear how the strength of these interactions will differ between the thiophene-6-azauridine parent and the extended-6-azauridines and how this non-covalent interaction will be affected upon excitation.

In order to gain an understanding of the excited state dynamics of molecular rotors is necessary to investigate both the ground and excited states. One method to theoretically investigate excited states is TD-DFT. This technique has been employed on a range of molecules that exhibit rotational behaviour.[13,14,15] For example, a TDDFT study into the S_1 potential energy curves of a distorted-BODIPY-based probe BV-1 was carried out in order to elucidate the nature of the twisting motion it undergoes in the excited state, after its fluorescence was shown experimentally to be sensitive to viscosity.[16] Similar calculations were also performed to prove the molecular rotor nature of a BODIPY derivative used to study intracellular viscosity changes in living cells.[17] This TD-DFT method will be used in the present work to investigate the molecular rotor mechanism of thiophene-6-azauridine.

Although there is little known about the excited state properties of thiophene-6-azauridine, there have been several studies into the excited state dynamics of 6-azauracil. A 2008 study by Ichimura *et al.* investigated the excited state characteristics of 6-azauracil in acetonitrile and compared the relaxation mechanism to that of the natural base uracil.[18] They used a combination of transient absorption spectroscopy and time-resolved thermal lensing, as well as using TD-DFT to calculate transition energies. The TD-DFT measurements indicated that

the first excited singlet state of 6-azauracil arises due to a HOMO-1 to LUMO transition with nπ* characteristics, while the S_2 state has predominantly ππ* character. This was also shown to be the case for uracil. Uracil has a T_1 state that lies 11000 cm^{-1} below the S_1 state, and 6-azauracil has a T_1 state 8900 cm^{-1} lower in energy than S_1, a smaller difference. The quantum yield for singlet to-triplet intersystem crossing of 6-azauricil was determined by transient absorption spectroscopy at 248 nm, using acetone as a sensitizer. It was found to be 1.00 ± 0.10, much higher than that of uracil (0.21 ± 0.02). When exciting 6-azauracil at 248 nm, a wavelength that results in excitation to the S_2 state, no fluorescence was observed. When exciting at 308 nm, a wavelength corresponding to excitation directly into the S_1 state, fluorescence emission was seen at 420 nm with a fluorescence quantum yield of 4.2 ± 0.4 x 10^{-3}. The quantum yield of intersystem crossing at this wavelength was found to be 0.93, slightly lower than the 248-nm value, with competition from a minor relaxation pathway due to internal conversion from S_1 to S_0. The slightly lower quantum yield for intersystem crossing when directly exciting the S_1 state was explained by vibrational relaxation form S_2 to S_1, resulting in a vibrationally hot S_1 state which has more excess energy than a directly excited S_1 state, thus promoting more rapid ISC.

To summarise, these measurements showed that the main relaxation pathway of excited state 6-azauracil is intersystem crossing from the singlet S_1 (nπ*) to the triplet state T_1 (ππ*). This differs from the major relaxation pathway of uracil which is internal conversion. A further study by Ichimura *et al.* in 2009 illustrated the importance of having a low lying nπ* state below the ππ*.[19] Alongside 6-azauracil they also investigated 8-azaadenine, 5-azacytosine and 8-azaguanine. Like 6-azauracil, 8-azaadenine has an nπ* below the first ππ*, both 5-azacytosine and 8-azaguanine have ππ* S_1 and S_2 states. It was found that like 6-azauracil, 8-azaadenine exhibits a high quantum yield of ISC whereas 5-azacytosine and 8-azaguanine do not. Similar measurements to those on 6-azauracil were carried out on 6-azauradine. The quantum yield for ISC at 248 nm was found to be the same as for 6-azauracil, 1.0, but excitation at 308 nm resulted in a lower yield, 0.73.[20] This illustrated the effect of the increased number of vibrational modes of the riboside. Unlike 6-azauricil where fluorescence was observed on exciting at 308nm, 6-azauridine was found to be non-fluorescent.

The measurements discussed this far were carried out in aprotic acetonitrile. It has been shown that for uracil in aprotic solvents the lowest singlet excited state has mostly nπ* character; in polar solvents this is not the case.[21] Instead this state lies higher than the ππ* state. Computational calculations showed that hydration lowers the energy of the ππ* state to below that of the nπ*.[22] Recently Liu *et al.* used transient absorption spectroscopy to investigate 6-

azauricil in water.[23] They concluded that on exciting at 266 nm S_2 ($\pi\pi^*$) decays to the S_1 ($n\pi^*$) within 0.3 ps, this then decays *via* ISC to T_1, the lifetime of which is longer than 1 ns. These results show that even in water, IC is negligible and ICS is the predominant mechanism for non-radiative decay. In this same study they also investigated uracil and they suggest an explanation for the fact IC crossing is seen in uracil but not 6-azauracil. Torsion about the C5=C6 double bond is known to be a key step in the uracil excited state decay process. 6-azauridine is unable to bend in the same manner as uracil due to the additional nitrogen and subsequently this decay pathway is blocked.

As thiophene-6-azauridine has a high quantum yield, intersystem crossing from the S_1 to T_1 cannot be a major excited state relaxation pathway. This suggests that unlike 6-azauradine which has a low lying $n\pi^*$ which facilitates this intersystem crossing, the thiophene-6-azauridine has a low lying $\pi\pi^*$ state.

In this Chapter, fluorescence lifetime measurements of the extended 6-aza-uridines are presented for the first time. Fluorescence lifetime measurements of the thiophene parent compound were made previously in various solvents.[2] These measurements were made using an LED as an excitation source, time-response and resolution were not reported. Fits of the resulting decays are not global and fitting has not been carried out from the peak of the decay. This suggests that short lifetime components would not be observed, even if present. All lifetimes were found to be mono-exponential, but due to the nature of the fit these results may not be conclusive. Re-measuring these lifetimes could therefore reveal decay components that were not previously resolved.

To aid in the understanding of the fluorescence decay characteristics of thiophene-6-aza-uridine and extended-6-aza-uridines, DFT and TDDFT calculations have been carried out. Firstly, the results for thiophene-6-azauridine will be presented, both experimental and computational. Secondly those results for the extended-6-azauridines will be considered.

4.2 Experimental

Thiophene-6-azauridine (parent) and the extended-thiophene-6-azauridines were received in powder form from the lab of Yitzhak Tor. Formal names and abbreviations used in this chapter are listed in Table 4-3. Samples were dissolved in either phosphate buffer (pH 7.4, Sigma-Aldrich), ethanol (>99.9%, Fisher Scientific), methanol (>99.9%, Fisher Scientific) or 1,4-dioxane (99.8%, extra dry, Acros Organics). Concentrations were in the range of ~ 1 x 10^{-5} M with exact concentrations determined from UV-vis measurements. All measurements were made at room temperature.

Formal Name	Abbreviations
5-(thiophen-2-yl)-6-aza uradine	Thiophene-6-azauridine or parent
5-(5-(4-fluorophenyl)thiophen-2-yl)-6-aza-uridine	1a , F derivative
5-(5-phenylthiophen-2-yl)-6-aza-uridine	1b, H derivative
5-(5-(4-methyl)phenyl)thiophen-2-yl)-6-aza-uridine	1c, Me derivative
5-(5-(4-methoxyphenyl)thiophen-2-yl)-6-aza-uridine	1d, OMe derivative
5-(5-(4-hydroxyphenyl)thiophen-2-yl)-6-aza-uridine	1e, OH derivative
5-(5-(4-(dimethylamino)phenyl)thiophen-2-yl)-6-aza-uridine	1f, NMe_2 derivative

Table 4-3. Formal names and abbreviations used in this thesis for thiophene-6-azauridine and its extended derivatives.

Computational calculations were performed using Gaussian 09.[24] Geometry optimisations were performed using crystal structures as the basis for starting geometries. Where necessary, modifications to crystal-derived geometries, such as modification of functional groups, were performed using Jmol or ArgusLab.[25,26] For example, for the parent compound, a crystal structure only exists for a brominated derivative. To establish ground-state torsional potentials, a relaxed scan was implemented in Gaussian, with bond angles varying by 10 or 15 degrees. Once ground state minima or transition states were identified these structures were then optimised with no dihedral angle constraint parameters. Ground-state frequency calculations were performed on these optimised structures, to confirm minima had been found, and from these calculations Gibbs free energies were obtained. For excited-state calculations, potential energies were used. This was done for several reasons: firstly to reduce computational time and cost; secondly, differences were not great between corrected and non-corrected energies and only comparisons between molecules were needed; and thirdly, because the thermochemistry analysis can produce errors when looking at molecules with hindered rotation. From these energies, conformational population distributions were calculated using the following equation:

$$\frac{N_{C1}}{N_{C2}} = \exp -(\frac{G_{C1} - G_{C2}}{RT})$$

Equation 4-1

where N_{C1} is the proportion of the population that exists as conformer 1; N_{C2} is the proportion of the population that exists as conformer 2; G_{C1} and G_{C2} are the free energies of the optimised

structures of conformer 1 and conformer 2, respectively; R is the gas constant and T is temperature in Kelvin.

The Eyring was used to calculate a rate constant for interconversion between conformers (k):

$$k = \frac{k_B T}{h} e^{-\frac{\Delta G^{\ddagger}}{RT}}$$

Equation 4-2

where k_B is Boltzmann's constant, h is Planck's constant and $\Delta G^{\ddagger}$ is the free energy of activation.

Time-resolved fluorescence measurements were made using the TCSPC setup described in Chapter 3. Excitation was performed at 310 nm for thiophene-6-azauridine, and at 360 nm for the extended 6-azauridines. Measurements were recorded at three emission wavelengths and decays were fitted globally. The emission wavelengths used are indicated alongside the corresponding lifetime results.

4.3 Results and Discussion

4.3.1 Thiophene-6-azauridine

4.3.1.1 Steady-state Fluorescence

Excitation and emission spectra of thiophene-6-azauridine have been collected in several different solvents. Those in dioxane, buffer and methanol are displayed in Figure 4-3. As has been reported previously, absorption and excitation spectra are largely unchanged between different solvents, with peak absorption consistently around 332 nm. Emission wavelength was found to be solvent-dependent. The smallest Stokes shift is seen in dioxane, where the emission maximum is at 415 nm, in methanol emission is at 433 nm and in phosphate buffer at 462 nm. The emission maximum in water is 455 nm, with a quantum yield of 0.20. In phosphate buffer the quantum yield was found to be 0.34, with a slightly longer emission wavelength. These differences are most likely due to difference in pH. Quantum yield measurements in methanol and dioxane were consistent with those previously reported, 0.5 and 0.8, respectively.

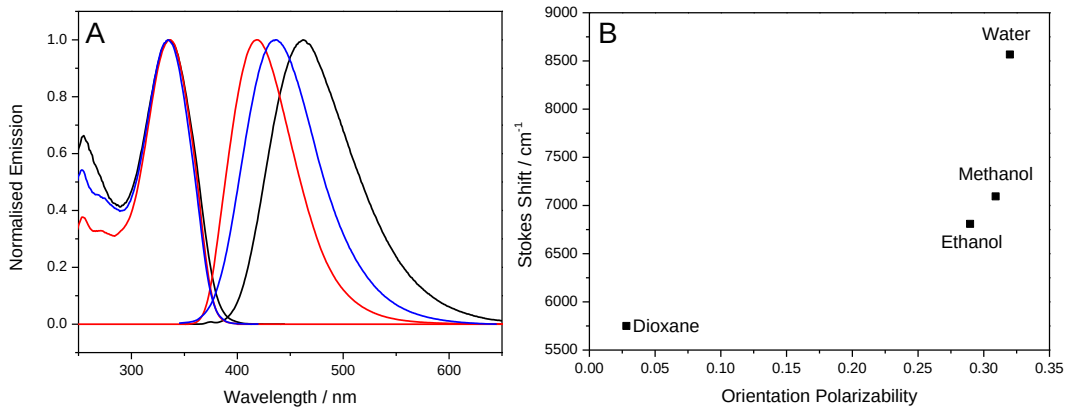


Figure 4-3. A) thiophene-6-azauridine excitation spectra, recorded at the peak emission wavelength, and emission spectra, excited at 330 nm, in buffer (black), dioxane (red) and methanol (blue). B) Lippert plot for thiophene-6-azauridine in the four solvents.

Previous polarity-dependence measurements by Tor *et al.* were carried out in dioxane/water mixtures and a linear relationship between Stokes-Shift and $E_T(30)$ values was found. The Lippert plot shown in Figure 4-3B is not linear. This could indicate that specific effects, most likely hydrogen bonding with water, not those due solely to polarity, are having an effect. It could also indicate that solvent is affecting the relative energy of excited states, for example the energy of a twisted internal charge-transfer state.

The thiophene-6-azauridine emission spectrum appears to be dependent on excitation wavelength. This was not reported in the previous study, in which spectra were recorded after exciting at the peak absorption wavelength, 332 nm. The extent to which wavelength-dependence is observed is also affected by the solvent. Emission spectra resulting from excitation at different wavelengths, in dioxane and phosphate buffer, respectively, are shown in Figure 4-4.

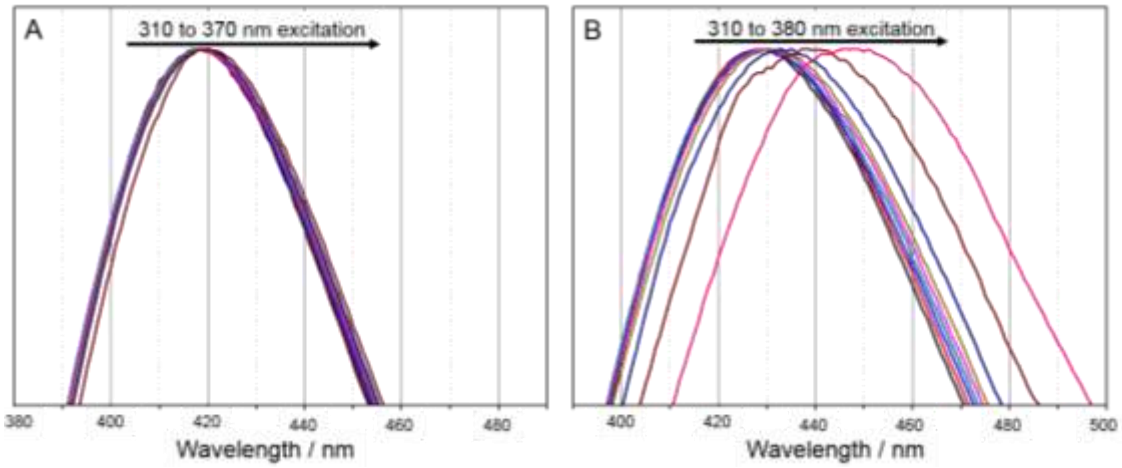


Figure 4-4. A) Normalised emission spectra of thiophene-6-azauridine in dioxane, excited at 10-nm increments between 300 and 370 nm. B) Normalised emission spectra of thiophene-6-azauridine in phosphate buffer, excited at 10-nm increments between 300 and 370 nm.

In dioxane, thiophene-6-azauridine was excited between 300 and 370 nm at 10-nm intervals. Although differences are small, the emission wavelength can be seen to increase with increasing excitation wavelength. A shift in peak emission wavelength of around 2 nm is found between exciting at 300 and 370 nm. In buffer, a greater shift in emission spectrum is seen when exciting at wavelengths between 300 and 380 nm. When excited at 300 nm, the emission maximum falls at 429 nm, when excited at 360 nm this increases to 435 nm and excitation at 380 nm results in an emission maximum of 446 nm. This wavelength-dependence indicates that there are multiple species present. As there is a larger difference in buffer than in dioxane, this could indicate that the emission spectra of the different species in buffer are more distinct than in dioxane.

4.3.1.2 Time-resolved Fluorescence

Fluorescence lifetimes were measured for thiophene-6-azauridine in various solvents. Previously fluorescence decays for the thiophene-6-azauridine parent have been reported to be mono-exponential. Fitting globally, across three emission wavelengths, to two lifetimes gave significantly better results than fits made using one lifetime. For a mono-exponential fit, χ^2 values were typically on the order of 2 to 5, significantly larger than the χ^2 of 1.09 achieved for a mono-exponential fit of a standard, coumarin 153. A significant improvement in quality of fit was obtained for a two-component fit, which gave χ^2 values typically less than 1.1. Examples of fluorescence decay data and fitted functions for thiophene-6-azauridine in methanol are displayed in Figure 4-5.

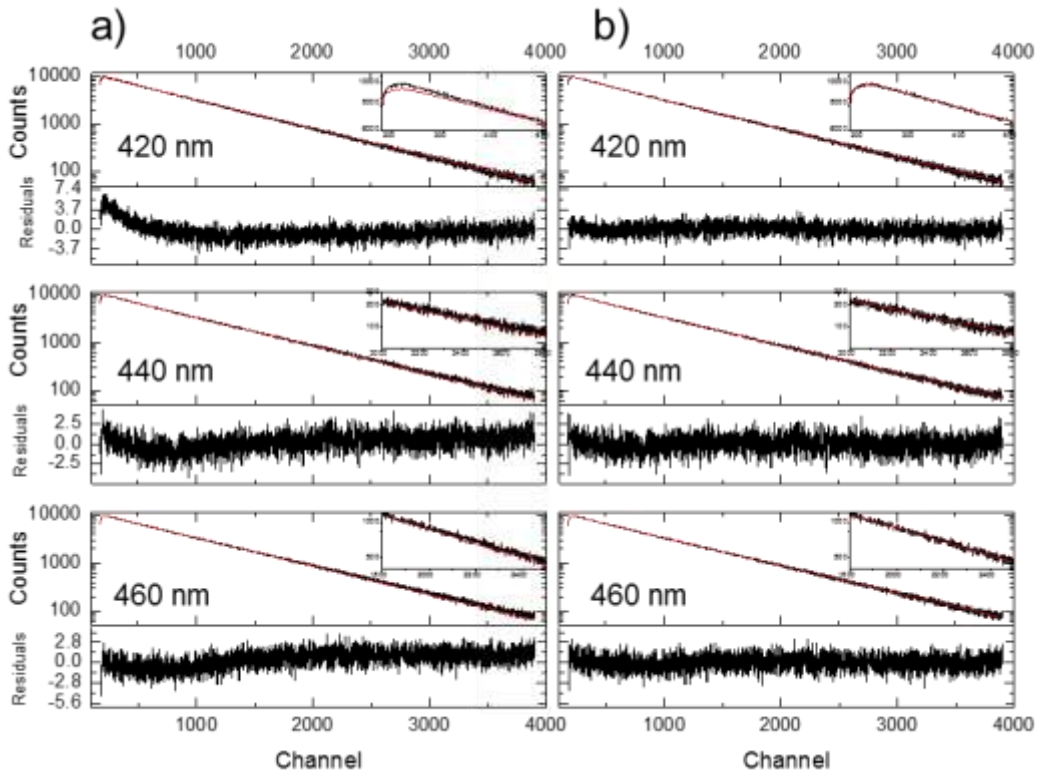


Figure 4-5. Fluorescence decay data (black) and fitted exponential functions (red), for thiophene-6-azauridine in methanol, excitation 310 nm. a) a one-component fit. b) a two-component fit. Below each decay are the corresponding residuals. Inserts contain an enlargement of a portion of the decay.

Figure 4-5a) compares a one-component fit and b) a two-component fit. The one component fit results in χ^2 values of 2.5 at 420 nm, 1.4 at 440 nm and 2.5 at 460 nm. This corresponds to a global χ^2 value of 1.9. The two component fit results in χ^2 values of 1.09 at 420 nm, 1.05 at 440 nm and 1.01 at 460 nm. This corresponds to a global χ^2 value of 1.05. By comparing the one and two component fits (most noticeable in the inserts) differences between residuals can be seen. This is most obvious at short times, where there is a large discrepancy between the fitted function and the experimental data, for the one component fit. It can also be seen that the residuals are not randomly distributed around zero for the one-component fits. This example was chosen as these data represent a case where a one-component fit could appear to be satisfactory, especially if only one emission wavelength were considered. In dioxane and buffer, the difference in goodness of fit between one and two components was significantly larger, and the choice between a one or two component fit significantly clearer.

The lifetimes and corresponding fractional amplitudes (A-factors) of thiophene-6-azauridine, excited at 310 nm, in dioxane, buffer, ethanol and methanol are presented in Table 4-4.

Solvent	τ_1 / ns	τ_2 / ns	A_1	A_2	$<\tau>$/ns	QY
Dioxane	0.1	6.7	0.28	0.72	4.8	0.80
Buffer	1.9	8.3	0.55	0.45	4.8	0.34
Ethanol	7.3	8.9	0.28	0.72	8.5	0.70
Methanol	6.0	9.4	0.15	0.84	8.8	0.50

Table 4-4. Fluorescence decay parameters of thiophene-6-azauridine parent in dioxane, buffered aqueous solution, ethanol and methanol, derived from global analysis of fluorescence decays recorded at 3 emission wavelengths. Lifetimes (τ_i) are global values; A factors (A_i) are quoted for the maximum emission wavelength; $<\tau>$ is the average lifetime (indicative of the quantum yield). Excitation performed at 310 nm. The quantum yield determined from steady-state fluorescence intensity is also shown. Error in τ_1 = 4%, Error in τ_2 = < 1%, Error in A_1 = 6%, Error in A_2 = < 1%

The biexponential decays exhibited by thiophene-6-azauridine in all solvents indicate the presence of two emitting species. The magnitude of the lifetimes and the fractional amplitudes of the two species vary depending on solvent. In buffer, the 1.9 ns component accounts for 55% of the population and the longer 8.3 ns component the remaining 45%. In dioxane τ_1 is significantly shorter than τ_2 and this short 0.1-ns component accounts for only 28% of the population. In methanol and ethanol the longer component also accounts for the majority of the population, 70 to 85%. Although both the magnitude of the lifetimes and the relative amplitude changes, two lifetimes are consistently observed for thiophene-6-azauridine in all solvents.

In all four solvents the relative proportion of the two species varies with emission wavelengths. This indicates that the two emitting species have differing emission spectra and is consistent with the steady state results from Figure 4-4. A-factors recorded at three emission wavelengths in buffer are presented in Table 4-5, dioxane Table 4-6, methanol Table 4-7 and ethanol Table 4-8.

λ_{em}/nm	A_1	SS_1	A_2	SS_2	$<\tau>$/ns
410	0.57	0.24	0.43	0.76	4.63
430	0.55	0.22	0.45	0.78	4.78
450	0.54	0.21	0.46	0.79	4.86

Table 4-5. Fluorescence decay parameters, fractional contributions to steady-state intensity (SS) and average lifetime ($<\tau>$) for thiophene-6-azauridine in buffer, at emission wavelength 410 nm, 430 nm and 450 nm. Lifetimes τ_1 = 1.9 ns and τ_2 = 8.3 ns

λ_{em}/nm	A_1	SS_1	A_2	SS_2	$<\tau>$/ns
400	0.34	0.01	0.66	0.99	4.46
420	0.28	0.008	0.72	>0.99	4.82
440	0.23	0.006	0.77	>0.99	5.17

Table 4-6. Fluorescence decay parameters, fractional contributions to steady-state intensity (SS) and average lifetime ($<\tau>$) for thiophene-6-azauridine in dioxane, at emission wavelength 400 nm, 420 nm and 440 nm. Lifetimes τ_1 = 0.1 ns and τ_2 = 6.7 ns

λ_{em}/nm	A_1	SS_1	A_2	SS_2	$<\tau>$/ns
420	0.26	0.18	0.74	0.82	8.48
440	0.15	0.10	0.84	0.90	8.84
460	0.12	0.08	0.88	0.92	8.96

Table 4-7. Fluorescence decay parameters, fractional contributions to steady-state intensity (SS) and average lifetime ($<\tau>$) for thiophene-6-azauridine in methanol, at emission wavelength 420 nm, 440 nm and 460 nm. Lifetimes τ_1 = 6.0 ns and τ_2 = 9.4 ns

λ_{em}/nm	A_1	SS_1	A_2	SS_2	$<\tau>$/ns
415	0.34	0.30	0.65	0.70	8.37
435	0.28	0.24	0.72	0.76	8.48
455	0.24	0.21	0.76	0.79	8.54

Table 4-8. Fluorescence decay parameters, fractional contributions to steady-state intensity (SS) and average lifetime ($<\tau>$) for thiophene-6-azauridine in ethanol, at emission wavelength 415 nm, 435 nm and 455 nm. Lifetimes τ_1 = 7.3 ns and τ_2 = 8.9 ns

In buffer, the 1.9 ns lifetime component accounts for 57% of the population at 410 nm whereas at 450 nm it accounts for 54%. In the other solvents more of a variation is observed. For example, in methanol at 420 nm the 6.02 ns lifetime component accounts for 26% of the population whereas at 460 nm is accounts for only 12%. The contribution of each species to the steady-state intensity is also displayed in Table 4-5 to Table 4-8. The short lifetime τ_1 contributes < 1% of the steady state spectrum in dioxane whereas in buffer, ethanol and methanol this shorter lifetime contributes between 8 to 30%. As the short lifetime component in dioxane contributes so little to the steady-state compared to the contribution of the second lifetime in buffer; this explains why a greater excitation-wavelength-dependence is observed for the emission spectrum in buffer compared to dioxane.

It was previously suggested that thiophene-6-azauridine could be a molecular rotor. One possible explanation for the two lifetimes it exhibits is that they are due to the existence of two conformers related by torsion about the thiophene-6azauridine inter-ring single bond.

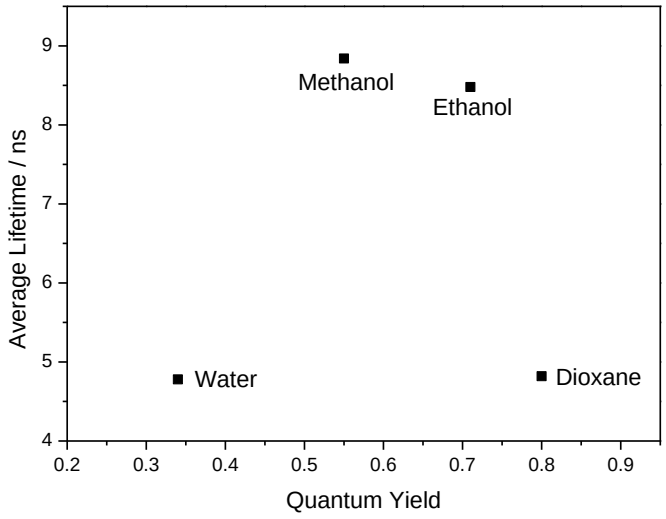


Figure 4-6. Average lifetime of the thiophene-6-azauridine plotted against the quantum yield in buffer, methanol, ethanol and dioxane.

Comparing average lifetime and quantum yield as illustrated in Figure 4-6 it can be seen that there is no linear relationship. In both buffer and dioxane the average lifetime, when calculated using the A factors from the peak emission wavelength, is 4.8 ns. Conversely, the quantum yield values differ by a large amount, 0.34 in buffer compared to 0.80 in dioxane. This discrepancy between average lifetime and quantum yield is indicative of the presence of dark states, i.e species whose lifetimes are too short to be detected in the present TCSPC measurements. If it is assumed that there are no dark states in dioxane, the solvent in which thiophene-6-azauridine has the highest quantum yield, the relative population of dark states, A_0, in the other three solvents can be calculated using the equation:

$$A_0 = 1 - \frac{\Phi_{rel}}{\langle \tau_{rel} \rangle}$$

Equation 4-3

where Φ_{rel} is the quantum yield relative to that in dioxane, $\langle \tau_{rel} \rangle$ is the average lifetime relative to that in dioxane and A_0 is the fractional population dark states.

	Φ_{rel}	$\langle \tau_{rel} \rangle$	A_0	A_1	A_2
Buffer	0.38	1.0	0.62	0.21	0.17
Ethanol	0.88	1.8	0.51	0.14	0.35
Methanol	0.62	1.8	0.65	0.06	0.29

Table 4-9. Quantum yields and average lifetimes relative to those of thiophene-6-azauridine in dioxane, A factors corrected for dark states are those at the peak emission wavelength. A_0 is the fractional population of dark states.

From Table 4-9 it can be seen that dark states make up a significant proportion, accounting for over half the population in all three solvents. In buffer the dark state accounts for 62% of the population, with the proportion of the emissive species fairly similar 21% A_1 and 17% A_2. In ethanol the dark state accounts for 51% and in methanol 65%. These values are calculated using the A factors measured at the peak emission wavelength, 330 nm in buffer, 335 nm in ethanol and 340 nm in methanol.

4.3.1.3 Computational Calculations

4.3.1.3.1 Ground State

In order to help understand the nature of the two emitting species measuring using TCSPC and the nature of the dark species, density functional theory (DFT) calculations have been carried out. As rotational isomerism has been identified as a possible source of conformers for thiophene-6-azauridine, DFT was used to determine the ground state geometry and torsional potential about the S(1)-C(10)-C(11)-N(4) bond in thiophene-6-azauridine. The B3LYP/6-311+G(d,p) functional and basis set was selected after benchmarking with a series of basis sets and functionals (Appendix 1). Calculations were initially carried out for molecules in the gas

phase. Two planar minimum-energy structures and a twisted transition state were found, as shown in Figure 4-7. Figure 4-7. From the given Gibbs free energies the population distribution of the two conformers and the lifetime for interconversion between them was calculated.

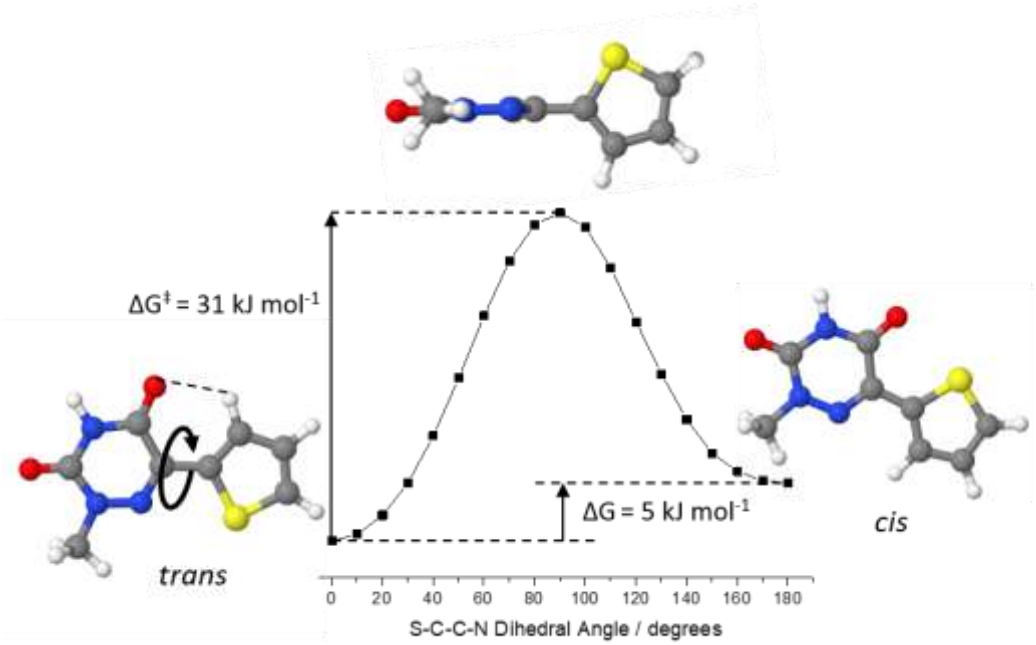


Figure 4-7. Global minimum structure (*trans* conformer), torsional isomer (local minimum) structure (*cis* conformer) and the calculated torsional potential. Transition state structure has the two ring planes perpendicular to each other. Calculations performed in the gas phase. Atoms: Carbon (grey), Hydrogen (white), Nitrogen (blue), Sulphur (yellow). Images created in jMol.

In the gas phase, it was found that energy of the planar conformers was such that they would both exist in the ground state, a ΔG equal to 5 kJ mol^{-1} corresponds to a population at room temperature in which 89% is *trans* conformer and 11% *cis* conformer. $\Delta G^{\ddagger}$ was found to be equal to 31 kJ mol^{-1} a barrier height that results in a timescale of interconversion for the *trans* to the *cis* conformer of 39 ns. It is therefore likely that both conformers will be present in solution and could be two of the species identified from lifetime measurements. However, the rate of interconversion means that the two conformers would not be revealed by NMR spectroscopy, and therefore would not have been identified in the initial characterisation by Tor *et al.*[4]

It is also important to consider the influence of solvent. Experimentally determined A-factors reveal that the emitting species exist in varying amounts depending on solvent. It is therefore possible that solvent will have an effect of the relative stability of these two conformers. The most accurate way to compute the effect of solvent is though the inclusion of explicit solvent molecules. This, however, can be very computationally demanding due to the number of

molecules required, or, if only a few solvent molecules are included can rely heavily on the starting arrangement of these solvent molecules. An alternative strategy, and the one used here, is to use an implicit model, such as the polarizable continuum model (PCM).[27] The PCM models solvent in terms of bulk properties such as dielectric constant. The molecule of interest is embedded in a solvent cavity that is created using a series of overlapping spheres, the size and position determined by the van der Waals radius of the atoms in the molecule. An iterative approach is used to describe how the charge distribution of the molecule affects the solvent cavity and then how that resulting change in solvent field affects the molecule.

The results of calculations performed using a PCM solvent model of dioxane (ε = 2.2099), water (ε = 78.3553), ethanol (ε = 24.852) and methanol (ε = 32.613) are presented in Table 4-10.

Solvent	ΔG / kJ mol^{-1}	*Trans* population	*Cis* population	$\Delta G^{\ddagger}$ /kJ mol^{-1}	Lifetime/ns for $C_1 \rightarrow C_2$
Gas Phase	5	0.89	0.11	30	39
Dioxane	4	0.82	0.18	29	16
Water	0.4	0.54	0.46	23	2
Ethanol	0.6	0.56	0.44	22	1
Methanol	0.6	0.56	0.44	23	2

Table 4-10. ΔG and corresponding *cis trans* conformer population distribution, height of barrier between *trans* and the transition state, $\Delta G^{\ddagger}$, and lifetime for interconversion from *trans* to *cis* calculated using DFT.

These results appear to indicate a strong solvent dependence. In dioxane the results are similar to those in the gas phase. ΔG equal to 4 kJ mol^{-1} results in a *trans* population that accounts for 82 % of the total. This is slightly lower than 89% of the population that is *trans* in the gas phase. In water, ethanol and methanol the energy of the two conformers are much more similar. Consequently, the *trans* population, accounts for just 54 -56%. $\Delta G^{\ddagger}$ is slightly lower in dioxane than in the gas phase, 29 kJ mol^{-1} compared to 30 kJ mol^{-1}, indicating that dioxane slightly stabilises the transition state. This effect is more pronounced in the more polar solvents indicating that they have a more stabilising effect.

It could appear that as there are two rotational isomers present and two emitting species determined from lifetime measurements (Table 4-4) that it would be possible to directly compare to the populations calculated from DFT to experimentally determined A factors. This however, does not account for the dark species.

4.3.1.3.2 Lifetime Calculation Comparison

In dioxane, experimentally it has been shown there are two emitting species, one with a lifetime of 0.1 ns that accounts for 28% of the population and a 6.7 ns component that accounts for 72%. This is consistent with the DFT calculations which predict two rotational conformers. The *trans* conformer accounts for the largest proportion, 82%, and a *cis* conformer accounts for the remaining 18%. As experimental A-factors and DFT calculations both describe population distribution these can be compared. As the *trans* conformer has been shown to account for 82% of the population and the 6.7 ns lifetime accounts for 72%, it is most likely that this 6.7 ns species is the *trans* conformer and the *cis* conformer is the 0.1 ns lifetime species.

Although only two emitting species have been detected in buffer, methanol and ethanol, by comparing lifetime and quantum yield results it has been shown that there are also dark species present. Consequently, the proportion of rotational conformers calculated using DFT cannot be directly compared to experimental A -factors. In all solvents it is likely that both rotational conformers exist but in buffer, methanol and ethanol there must be at least one additional component, most likely a tautomer. It is possible that tautomers also exist in dioxane but that the lifetimes are too similar to be reliably resolved or that the tautomer is only present in protic solvent.

In dioxane the two lifetimes are significantly different, 0.1 and 6.7 ns. In the other three solvents both the lifetimes are at least several nanoseconds. It is possible that the *cis* conformer responsible for the 0.1 ns component in dioxane is too short to detect in the more polar solvents and so both the lifetimes observed in these solvents are due to the *trans*-rotational isomer and a *trans*-tautomer. This is represented schematically for thiophene-6-azauridine in dioxane and methanol in Figure 4-8.

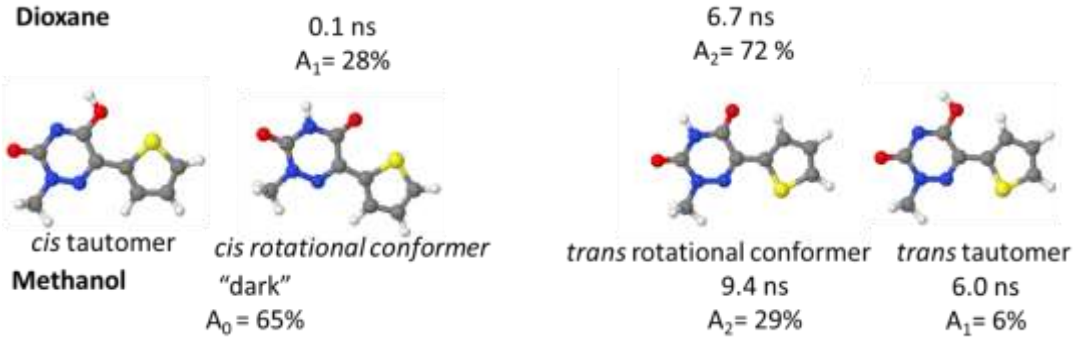


Figure 4-8. Thiophene-6-azauridine *cis* and *trans* conformers alongside possible tautomers of each. Diagram annotated with fluorescence lifetimes and A factors recorded in dioxane and methanol.

A comparison can be made between the A factors if the proportion of dark species is also included. A_0 is compared to the *cis* conformer population and $A_1 + A_2$ is compared to the *trans* population. A_0 was found experimentally to account for 65% total population in methanol. Computationally the *cis* conformer was only found to account for 44% of the population. There are several possibilities that can explain the differences. Firstly the A-factors are determined from exciting at just one wavelength and recording emission also at one wavelength. For example, Table 4-7 illustrates how large an effect this can have with the A factors varying from 26 to 12% in methanol when recording at an emission wavelength of 420 nm compared to 460 nm. The computationally determined populations on the other hand describe the total population.

The *trans* conformer is stabilised by a hydrogen bond between the 6-aza oxygen and a thiophene hydrogen (Figure 4-7). It is possible that competing hydrogen bonding from buffer could affect the degree to which this stabilisation occurs. The polarisable continuum model of solvation used in these calculations does not account for specific interaction. The hydrogen bonding effect of buffer will therefore not be accounted for. It is possible that if specific solvation was included, without the stabilising effect of this hydrogen bond, the *cis* thiophene-6-azauridine conformer would be found to be more populated.

4.3.1.4 Excited State DFT Calculations

Fluorescence lifetime measurements and ground-state DFT calculations indicate that the thiophene-6-azauridine exists as two rotational conformers, both present in the ground state. If the thiophene-6-azauridine behaves as a molecular rotor, as previously suggested, it could be expected that in the excited state a twisted conformation would be the lowest energy. It is also possible that in the excited state these planar species could interconvert with a lower barrier to rotation, or just one species will dominate in the excited state. The excited state potential energy as a function of inter-ring torsion was therefore investigated. For comparison, the furan-6-azauridine was also investigated. Previously the furan-6-azauridine fluorescent base analogue was shown to display a sensitivity to viscosity not exhibited by the thiophene-6-azauridine.[2] Both structures are displayed in Figure 4-9.

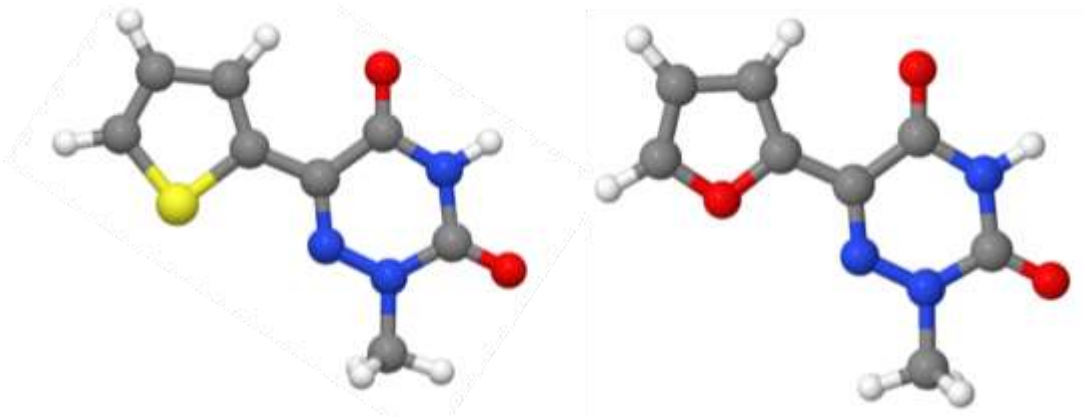


Figure 4-9. *Trans* structure of thiophene-6-azauridine (left) and *trans* structure of furan-6-azauridine (right). Atoms: Carbon (grey), Hydrogen (white), Nitrogen (blue), Sulphur (yellow). Images created in jMol.

Thiophene-6-azauridine was optimised in the ground state with the S(1)-C(10)-C(11)-N(4) dihedral fixed at angles between 0 and 180 degrees. The functional CAM-B3LYP was used as it has been shown to most accurately calculate energies of twisted excited states of similar models, whereas B3LYP has been shown to underestimate these.[28,29,14] With the S(1)-C(10)-C(11)-N(4) dihedral angle constrained, thiophene-6-azauridine was vertically excited to the first excited singlet state. This was followed by an excited state optimisation performed with the dihedral still fixed. These calculations were carried out in the gas phase. The ground state and optimised excited state (S_1) energy is plotted in Figure 4-10.

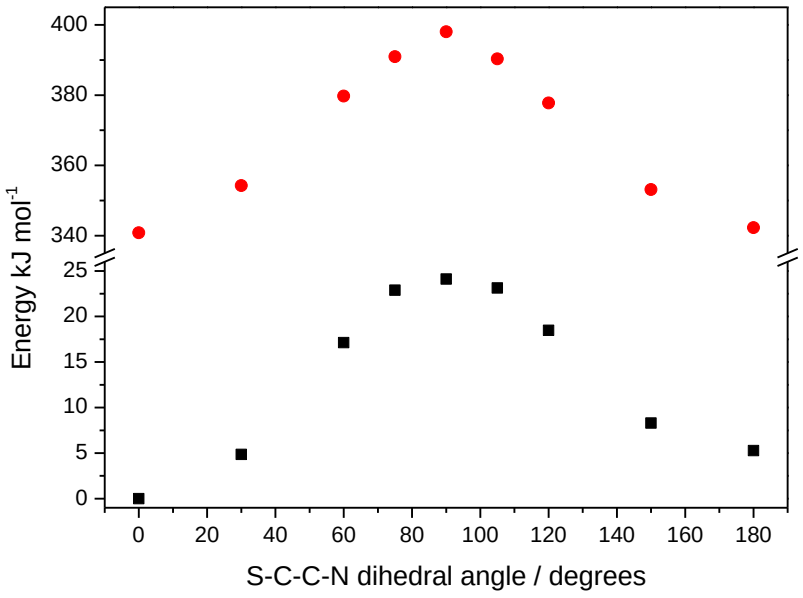


Figure 4-10. Potential energy of inter-ring torsion for thiophene-6-azauridine. Ground state (S_0) energy (black), excited state S_1 energy (red).

It can be seen from Figure 4-10 that the shape of the ground and excited state potential energy curves are similar; both have local minima when thiophene-6-azauridine is in a planar conformation and have a transition state at 90 degrees. The relative proportion of the two conformers in the ground and excited state differs. In the ground state the equilibrium ratio of *trans* to *cis*, using this functional, is 89:11. In the excited state it is found to be 64:36. The barrier to rotation is found to be much larger in the excited state than in the ground state, 57 kJ mol^{-1} compared to 24 kJ mol^{-1}. A barrier of 57 kJ mol^{-1} corresponds to a time for interconversion of 2 ms, several orders of magnitude longer than the fluorescence lifetime of the thiophene-6-azauridine. Consequently, it is predicted that there is no excited-state interconversion and emission is dictated by the population of the ground state species.

The ground-state and excited-state potential energy curves of furan-6-aza-uridine, calculated in an identical manner, are plotted in Figure 4-11.

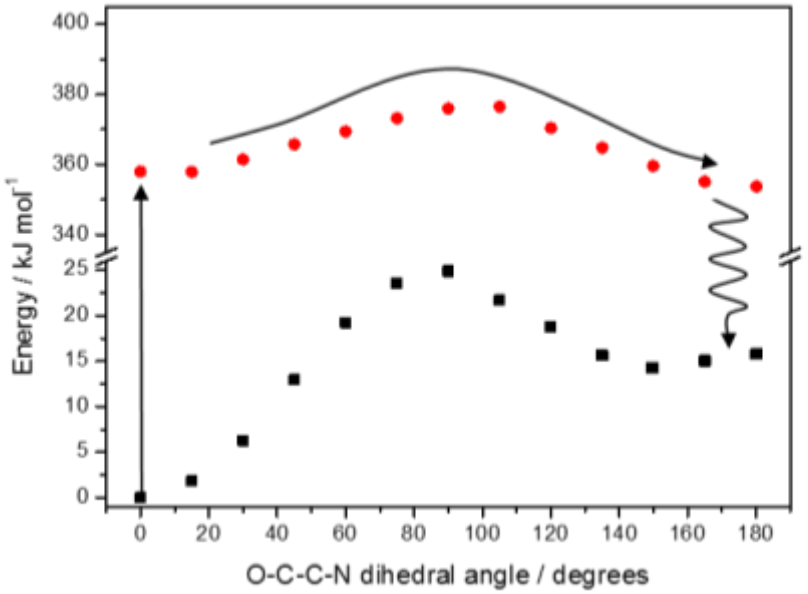


Figure 4-11. Torsional potential energy about O-C-C-N dihedral of furan-thiophene-6-azauridine. Ground state energy (black), S_1 energy (red). Arrows represent excitation, excited state interconversion *trans* to *cis*, followed by internal conversion to the ground state.

Differences between the furan-6-aza-uridine and the thiophene-6-aza-uridine can be elucidated by comparing Figure 4-10 and Figure 4-11. Firstly, differences in the ground state can be seen. The large difference in free energy, ΔG equal to 16 kJ mol^{-1}, between the two conformers for the furan compound (at room temperature) means that it will exist entirely as one isomer, the *trans*. The *cis* conformer is unfavourable due to repulsion between lone pair electrons on the oxygen atoms. The barrier between the planar and twisted state is of a similar magnitude to that observed for the thiophene-6-aza-uridine, 25 kJ mol^{-1}.

In the excited state, the *cis* isomer is shown to be slightly lower in energy than the *trans* conformer (the lower energy ground state conformer). The difference between excited-state energy and ground state is also at a minimum when in the *cis* conformation. The energy barrier to rotation in the excited state of the furan-6-azauridine is lower in the excited state than in the ground state, only 18 kJ mol^{-1}. This is substantially lower than the excited state barrier of 57 kJ mol^{-1} observed for the thiophene-6-azauridine. $\Delta G^{\ddagger}$ equal to 18 kJ mol^{-1} corresponds to a rate of interconversion of 0.2 ns. This suggest that the furan-6-azauridine could undergo excited state interconversion, from the *trans* to the *cis* state, followed by internal conversion to the ground state. No emission at a wavelength to the red of the dominant fluorescence has been reported, suggesting that the excited *cis* conformer undergoes rapid non-radiative decay.

Torsional motion appears to be an important pathway to excited state relaxation for the furan compound. For the thiophene-6-azauridine, two rotamers are present in both the ground and excited states and the barrier to interconversion is twice as high in the excited state. These differences explain the different response to viscosity between the furan and thiophene compounds. An increase in quantum yield is observed for the furan-compound with increasing viscosity due to its torsional relaxation pathway becoming hindered. For the thiophene-6-azauridine, both species are already present in the ground state and not interconverting in the excited state and thus an increase in viscosity will have less effect. This also explains the difference in quantum yield between the two. The quantum yield of the thiophene compound is higher than that of the furan compound in all solvents. This could be because the thiophene-6-azauridine does not have access to a torsional relaxation mechanism. The quantum yield is most different in polar solvents; in methanol thiophene-6-azauridine has a quantum yield of 0.5, whereas furan-6-azauridine has one of 0.04. This suggests that polar solvents have a stabilising effect on the twisted intermediate, lowering the barrier to rotation and making this pathway more favourable. To confirm this, it would be advantageous to perform excited state calculations in the presence of solvent.

4.3.2 Extended-6-aza-uridines

4.3.2.1 Steady-state Spectroscopy

The excitation and emission spectra of the extended-6-azauridines in dioxane and buffer are displayed in Figure 4-12.

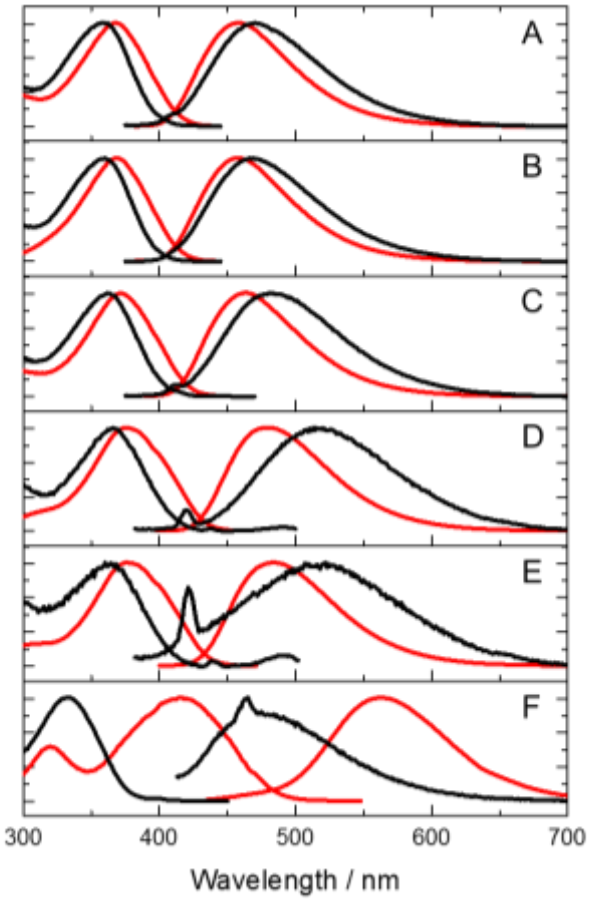


Figure 4-12. Excitation and emission spectra of the extended 6-aza-uridines in buffer (black) and dioxane (red). A) 1a Y=F, B) 1b Y=H, C) 1c Y=Me D) 1d Y=OMe E) 1e Y=OH F) 1f NMe_2

As previously reported and presented in Section 4.1, increasing the electron-rich character of the Y substituent leads to a red-shift in the absorption and emission spectra. 1a (Y=F) has the least red-shifted absorption and emission and 1f (Y= NMe_2) the most. For compounds 1a-1e in buffer the excitation spectrum is less red-shifted than that in dioxane and the emission spectrum is more red-shifted. 1f behaves differently; in buffer both the excitation and emission spectra are blue-shifted compared to those in dioxane. 1f contains a NMe_2 group which is often associated with a charge transfer transition.

4.3.2.2 Time-resolved Fluorescence

The lifetimes and corresponding fractional amplitudes (A-factors) of the extended 6-aza-uridines in dioxane are presented in Table 4-11 and those in buffer are presented in Table 4-12.

The extended-6-aza uridines in dioxane can be divided approximately into groups based upon these lifetime results. The first grouping contains 1a, 1b and 1c. Each has three lifetime components, a short lifetime of around 200 ps with an A factor of around 0.33, a lifetime of 1

to 2 ns with an A factor of around 0.45 and a component with a longer lifetime of around 4.5 ns and an A factor of approximately 0.22. The second group contains 1d and 1e, both exhibit two lifetimes, one component has a lifetime of ~0.3 ns with an A factor of 0.32 and a second has a longer lifetime of 4.25 ns and an A factor of 0.68. 1f is similar to 1d and 1e in terms of number of lifetimes but its longer lifetime is significantly shorter.

	Emission Wavelength /nm	τ_1 / ns	τ_2 / ns	τ_3 / ns	A_1	A_2	A_3	$<\tau>$ /ns	Quantum Yield
1a (F)	465	0.2	1.2	4.4	0.32	0.49	0.18	1.5	0.34
	485				0.32	0.44	0.24	1.7	
	505				0.32	0.39	0.29	1.8	
1b(H)	440	0.2	1.1	4.1	0.36	0.54	0.10	1.0	0.24
	480				0.33	0.45	0.22	1.4	
	500				0.33	0.41	0.26	1.6	
1c(Me)	470	0.2	1.8	4.8	0.34	0.51	0.15	1.7	0.34
	490				0.33	0.47	0.20	1.9	
	510				0.31	0.44	0.25	2.0	
1d(MeO)	470	0.3	-	4.2	0.37	-	0.63	2.8	0.51
	490				0.32	-	0.68	3.0	
	510				0.28	-	0.73	3.1	
1e(HO)	470	0.3	-	4.3	0.38	-	0.62	2.8	0.52
	490				0.33	-	0.67	3.0	
	510				0.29	-	0.71	3.2	
1f(Me_2N)	545	0.4	-	2.6	0.37	-	0.63	1.8	0.12
	565				0.37	-	0.63	1.8	
	585				0.39	-	0.61	1.8	

Table 4-11. 1a -1f decay parameters, in dioxane, globally fitted lifetimes, A factors displayed at three wavelengths. Excitation at 362 nm. Error in τ_1 = 6 %, Error in τ_2 = 3 % Error in τ_3 = <1 %. Error in A_1 = 3%, Error in A_2 = 3 % Error in A_3 = 1 %.

In phosphate buffer all lifetimes are shorter than those in dioxane. 1a, 1b, and 1c have three lifetimes in both solvents. The biggest difference is the long lifetime component, which decreases from 4.5 ns in dioxane to around 1 or 2 ns. The A-factor of this component increases from ~22% to ~33%. 1d in buffer has two lifetimes, the same as in dioxane, both these lifetimes are significantly shorter than those in dioxane. The majority, 72%, of the population is a component with a 70 ps lifetime. The quantum yields of 1e and 1f are so low in buffer that lifetimes could not reliably be recorded.

	Emission Wavelength /nm	τ_1 / ns	τ_2 / ns	τ_3 / ns	A_1	A_2	A_3	$\langle\tau\rangle$ /ns	Quantum Yield
1a (F)	450	0.1	0.7	1.8	0.38	0.41	0.45	0.78	0.1
	480				0.41	0.25	0.34	0.83	
	500				0.45	0.21	0.35	0.82	
1b(H)	450	0.1	0.7	2.0	0.36	0.37	0.28	0.84	0.12
	480				0.40	0.27	0.33	0.88	
	500				0.47	0.21	0.32	0.83	
1c(Me)	480	0.1	0.4	1.0	0.42	0.25	0.32	0.49	0.05
	500				0.40	0.30	0.31	0.49	
	520				0.39	0.32	0.28	0.48	
1d(MeO)	505	0.07	0.2	-	0.74	026	-	0.10	0.008
	525				0.72	0.28	-	0.10	
	545				0.70	0.30	-	0.10	

Table 4-12. 1a -1f decay parameters in phosphate buffer, globally fitted lifetimes, A-factors are those from the peak emission wavelength. Excitation at 362 nm. Quantum yields of 1e and 1f too low for lifetimes to be reliable. Error in τ_1 = 6 %, Error in τ_2 = 3 % Error in τ_3 = <1 %. Error in A_1 = 3%, Error in A_2 = 2 % Error in A_3 = 1 %.

The average lifetime and quantum yield of the extended-6-azauridines in buffer and dioxane follow a straight line trend except for 1f (Y=NMe_2) which is an outlier, as shown in Figure 4-13. When comparing the relative quantum yield and relative lifetimes of 1d to those of the other extended 6-aza-uridines it can be seen that 1d (Y=MeO) in buffer is also an outlier. This indicates that 1f in dioxane and 1d in buffer have dark states.

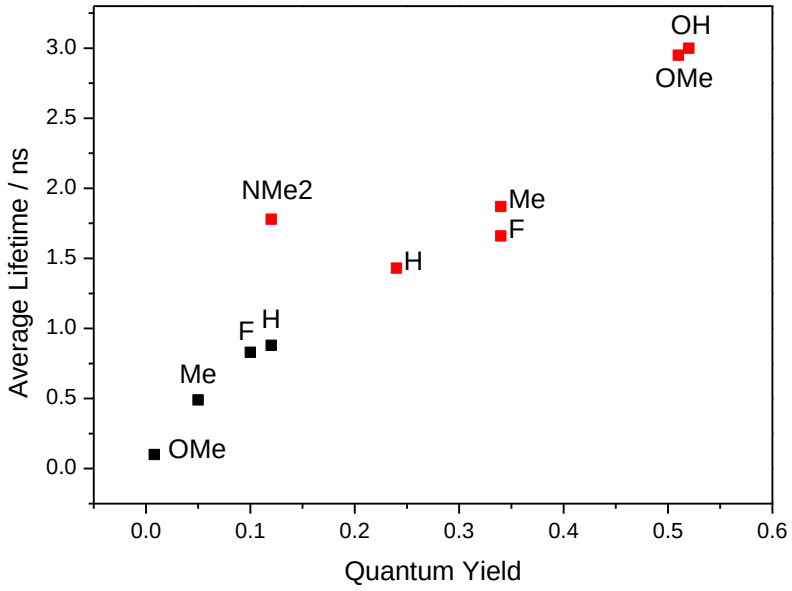


Figure 4-13. Average lifetime of the extended 6-aza-uridines in buffer (black) and dioxane (red) plotted against quantum yield.

4.3.2.3 Computational Calculations

It is possible that the lifetimes exhibited by these extended compounds are due to ground state conformers, similar to those seen for the parent compound. DFT calculations similar to those performed on thiophene-6azauridine have therefore been carried out for the extended-6-aza-uridines. Initially rotation with respect to the S(1)-C(10)-C(11)-N(4) dihedral, the same dihedral as in thiophene-6-azauridine, was investigated. Figure 4-14 illustrates this rotation. Just as for thiophene-6-azauridine, energies have been calculated using the functional basis set pair CAM-B3LYP/6-311+g(d,p).

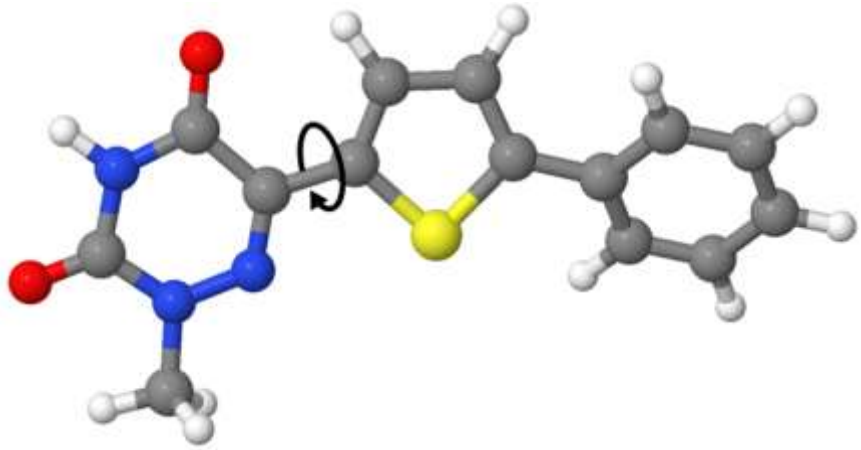


Figure 4-14. Ground state optimised structure of *trans* conformer of 1b (Y=H). Arrow indicates the torsional axis corresponding to the S(1)-C(10)-C(11)-N(4) dihedral angle.

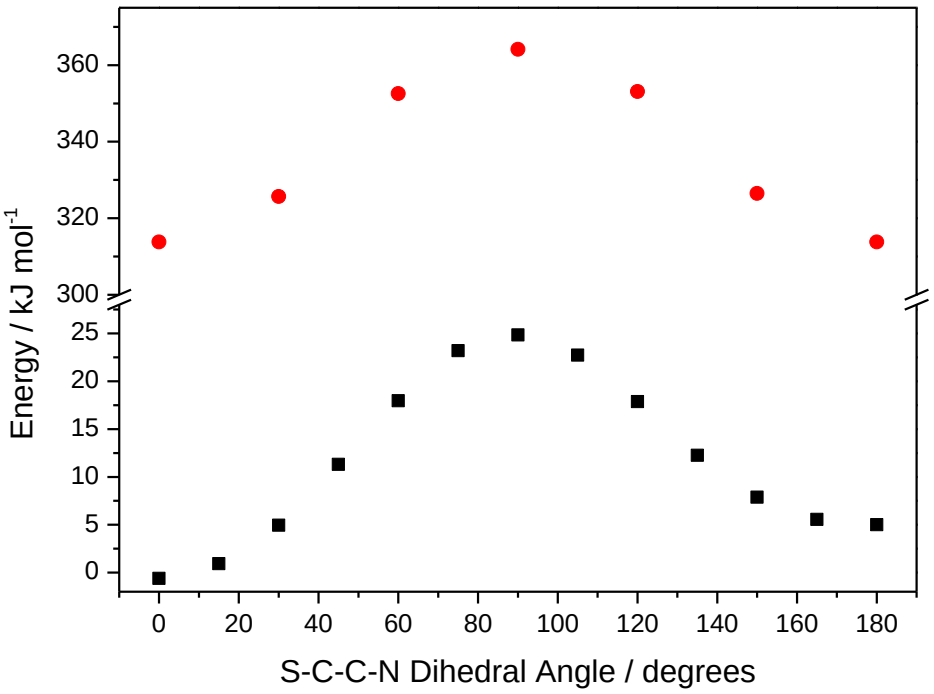


Figure 4-15. Potential energy of rotation around S-C-C-N bond for 1b extended-6-aza-uridine. Ground state energy (black), S1 energy (red).

The ground state potential energy as a function of angle about the S(1)-C(10)-C(11)-N(4) dihedral for 1b is illustrated in Figure 4-15. Potential energy curves look similar for all six of the extended 6-azauridines. ΔG, $\Delta G^{\ddagger}$ and the resulting conformer population distributions, and interconversion times, for all six of the extended-6-azauridines are summarised in Table 4-13.

	ΔG / kJ mol^{-1}	*Trans* population	*cis* population	$\Delta G^{\ddagger}$ /kJ mol^{-1}	Lifetime/ns for $C_1 \rightarrow C_2$
parent	5	0.89	0.11	24	5
1a (F)	5	0.90	0.10	26	6
1b(H)	6	0.91	0.09	25	5
1c(Me)	6	0.91	0.09	26	5
1d(MeO)	6	0.91	0.09	26	5
1e(HO)	6	0.92	0.08	26	6
1f(Me_2N)	8	0.96	0.04	27	7

Table 4-13. ΔG and corresponding *cis trans* conformer population distribution, height of barrier between *trans* and the *trans*ition state, ΔG‡, and lifetime for interconversion from *trans* to *cis* calculated using DFT for extended 6-aza-uridine 1a-1f and the parent for reference.

The ground state conformer population distribution is similar for all the extended 6-aza-uridines and also similar to those results obtained for the thiophene-6-azauridine. The parent exists as 89% the *trans* conformation, 1a to 1e very similar with 90-92% of their populations existing as the *trans* conformation. Compared to the parent, 1f shows the greatest difference, the *trans* conformer accounting for 96% of the total population compared to ~91% for the other extended-6-azauridines. The height of the barrier between the planar and twisted conformations is consistently higher by 2 or 3 kJ mol^{-1} compared to $\Delta G^{\ddagger}$ for thiophene-6-azauridine. This results in a lifetime for interconversion of the same or slightly longer, 5-7 ns. This shows that the ground state behaviour of the extended-6-azauridines, at least as it pertains to this torsional coordinate, is the same as the parent.

The optimised excited state (S_1) energy of 1b as a function of the S(1)-C(10)-C(11)-N(4) dihedral angle has also been plotted in Figure 4-15. The corresponding results for all the extended-6-azauridines are displayed in Table 4-14.

	ΔG /kJ mol^{-1}	*trans* population	*cis* population	$\Delta G^{\ddagger}$ /kJ mol^{-1}	Lifetime/μs for $C_1 \rightarrow C_2$
parent	1.4	0.64	0.36	57	1739
1a (F)	0.2	0.52	0.48	53	308
1b(H)	0.6	0.55	0.44	51	140
1c(Me)	0.2	0.52	0.48	49	51
1d(MeO)	-0.9	0.41	0.59	45	11
1e(HO)	-0.2	0.48	0.52	45	11
1f(Me_2N)	-0.6	0.44	0.56	36	0.3

Table 4-14. Excited state ΔG and corresponding *cis* and *trans* conformer population distribution, height of barrier between *trans* and the *trans*ition state, ΔG‡, and lifetime for interconversion from *trans* to *cis* calculated using DFT. for extended 6-aza-uridine 1a-1f and the parent for reference.

In the excited state there are greater differences between each of the extended-6-aza-uridines, and between the extended compounds and the parent. The parent population is 64% *trans* 36% *cis*. The extended compounds range from 55- 41% being *trans*. 1a, 1b and 1c favour the *trans* conformation whereas 1d, 1e and 1f favour the *cis*. The barrier height is also substantially different, both compared to each other and compared to the parent. $\Delta G^{\ddagger}$ is 57 kJ mol^{-1} for thiophen-6-azauridine, 53 kJ mol^{-1} for 1a reducing in height in the order 1a>1b>1c>1d~1e>1f. Despite the lower barrier to excited state rotation, rotation it is still on the timescale of μs and the barrier is sufficiently high such that excited state interconversion will not occur. As with the parent, the ground state population therefore dictates the *cis trans* population distribution.

In the ground state, the lowest energy conformation occurs for all extended-6-azauridines when the phenyl ring is at a ~30° angle to the rest of the molecule. This is illustrated in Figure 4-16 which displays the ground state optimised structure of 1b.

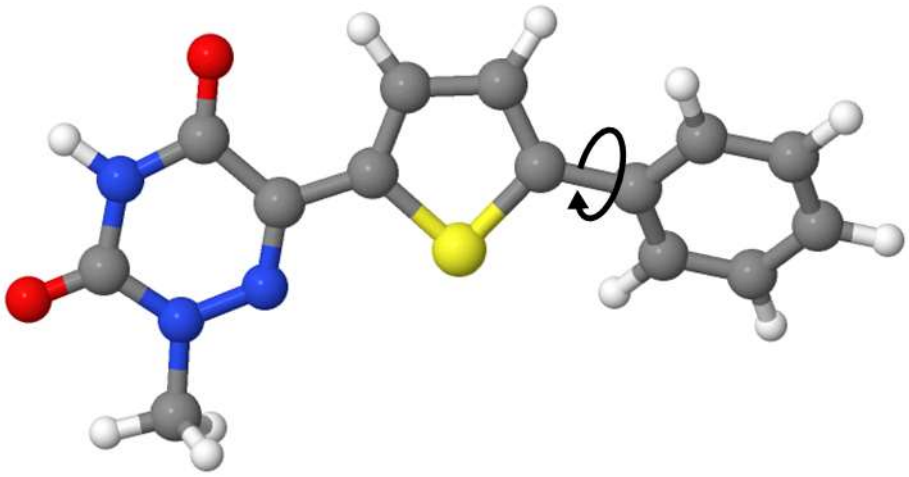


Figure 4-16. Optimised ground state structure of 1b. Arrow represents bond around which phenyl ring rotation refers to.

The presence of a second inter-ring single bond between the thiophene and the phenyl substituent means rotation of the phenyl ring is possible. Rotation about this S-C-C-C bond, indicated with an arrow on Figure 4-16, has also been investigated. The potential energy curve due to rotation about this bond, for 1b, has been plotted in Figure 4-17. In the ground state, an angle of 30 degrees between the thiophene and phenyl ring is most favourable, in the excited state it is a planar structure that is favoured. The potential is symmetrical, with energy the same if the bond is rotated by $+30^{o}$ or -30^{o}.

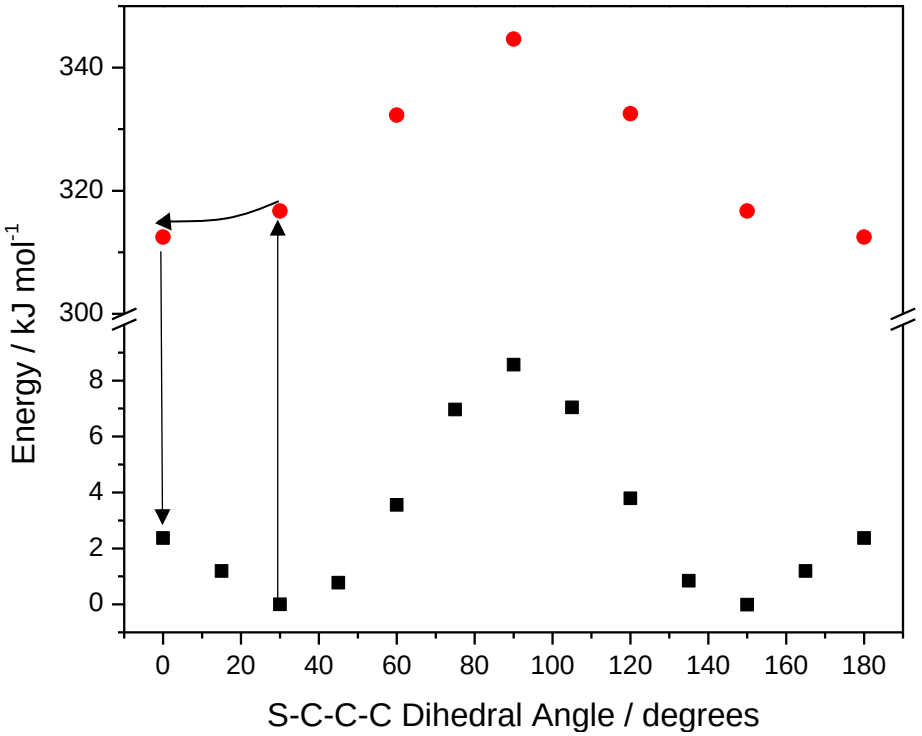


Figure 4-17. Potential energy of rotation around second ring (S-C-C-C bond) for 1b extended-6-aza-uridine. Ground state energy (black), S1 energy (red). Vertical arrows illustrate excitation and emission, curved arrow indicates torsion.

The difference in ground state energy between the 30 degree minimum and the 0 degree and 90 degree transition states for all the extended-6-azauridines are displayed in Table 4-15.

	$\Delta G^{\ddagger}$ 0 to 30 degrees /kJ mol^{-1}	$\Delta G^{\ddagger}$ 30 to 90 degrees /kJ mol^{-1}	Lifetime/ps for $C_{30} \rightarrow C_{150}$ *via* 0° barrier	Lifetime/ps for $C_{30} \rightarrow C_{150}$ *via* 90° barrier
1a (F)	2.5	8.0	0.4	4
1b(H)	2.4	8.6	0.4	5
1c(Me)	2.4	8.8	0.4	6
1d(MeO)	2.3	8.6	0.4	5
1e(HO)	2.3	8.3	0.4	5
1f(Me_2N)	1.7	9.4	0.3	7

Table 4-15. Ground state second ring rotation of extended-6-aza-uradines. $\Delta G^{\ddagger}$ 30 to 0 is the ground state barrier height between the 30 minimum and the planar transition state. $\Delta G^{\ddagger}$ 30 to 90 is the ground state barrier height between the 30 minimum and the twisted transition state corresponding lifetimes for interconversion also shown.

The barrier heights displayed in Table 4-15 are similar for all extended-6-azauridines with 1f having a slightly lower barrier between the identical 30-degree and 150-degree structures, when going *via* a planar transition state than when going *via* a twisted transition state. The timescale for interconversion between the two minima is on the order of picoseconds,

indicating that at room temperature these two identical structures would be rapidly interconverting.

In the excited state, the minimum energy structures are planar, with a single barrier at 90 degrees. $\Delta G^{\ddagger}$ for the excited state barrier of all the extended 6-azauridines is given in Table 4-16.

	$\Delta G^{\ddagger}$ /kJ mol^{-1}	Lifetime/ns for $C_0 \rightarrow C_{180}$
1a (F)	32	71
1b(H)	32	71
1c(Me)	35	215
1d(MeO)	40	1458
1e(HO)	38	659
1f(Me_2N)	53	377000

Table 4-16. Excited state torsional barrier between, $\Delta G^{\ddagger}$, for extended 6-azauridines and corresponding time for interconversion.

The nature of the Y substituent has a significant effect on the barrier height. For 1a and 1b $\Delta G^{\ddagger}$ is equal to 32 kJ mol^{-1} this increases to 53 kJ mol^{-1} for 1f. This indicates that, following vertical excitation to the 30 degree structure, the phenyl ring will rotate in the excited state directly to a planar conformation, not going *via* a twisted intermediate. This pathway is illustrated on Figure 4-17.

It has been shown experimentally that there are no dark states, aside from in 1f (NMe) and 1d (MeO) in buffer. DFT calculations have shown similar population distributions, with *cis* and *trans* populations dictated by ground state distributions, as the thiophene-parent. Consequently, the lifetimes can be interpreted as follows. In dioxane, 1d (MeO) and 1e (OH) display two lifetimes which are similar to those of the parent. These lifetimes can be assigned to the same conformers as the thiophene-parent. τ_1, the ~ 0.3 ns component is the *cis* conformer and the longer lifetime τ_3 (~ 4.3 ns) is the *trans* conformer. For 1a (F), 1b (H) and 1c (Me) τ_1 and τ_3 are similar to the two lifetimes observed for 1d and 1e. It is likely that similarly the short lifetime corresponds to the *cis* conformer and the long lifetime to the *trans* conformer. Comparing the A-factors, the population of τ_1, the *cis* conformer, is around 33% for all five compounds, suggesting that the *trans* population has separated into two separate population for 1a, 1b and 1c, most likely two *trans* isomers. For 1a, 1b and 1c in buffer although lifetimes are shorter, A factors are similar. A_1 accounts for 40% of the population and $A_2 + A_3$ is equal to 60%. τ_1 can be assigned to the *cis* isomer and τ_2 and τ_3 to the *trans*. In buffer, 1d shows evidence of dark states; the population of these accounts for ~40% suggesting the *cis*

conformer can be attributed to this dark species. The lifetimes of the *trans* species are also greatly reduced in buffer.

4.3.3 Comparison to Crystal Structures

Initial thiophene-6-azuridine starting geometries were taken from the crystal structures deposited by Tor *et al*. The crystal structure used was that of 5-(5-bromothiophen-2-yl)-6-aza-uridine. This differs from that of the thiophene-6azauridine nucleoside by having a bromine in place of a hydrogen as shown in Figure 4-18a.

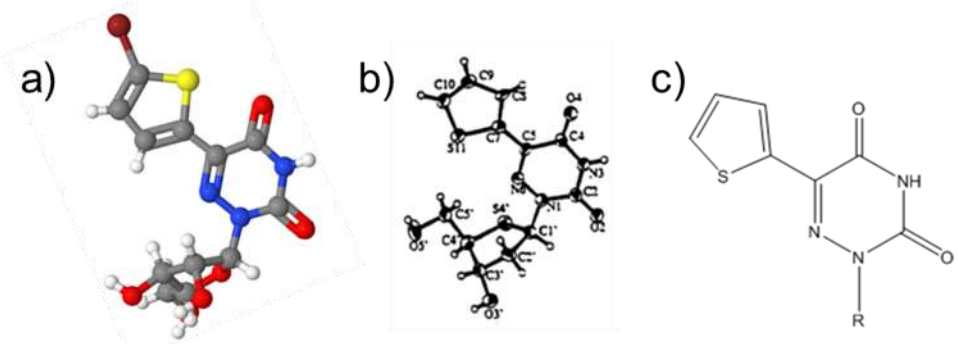


Figure 4-18. a) Crystal structure recreated from reference 4. b) Figure replicated from reference 27 and c) structural representation of crystal structure in b) shown for clarity.[30]

The most obvious structural difference, aside from the bromine, is that the predominant conformer calculated using DFT is the *trans* whereas the crystal structure is that of the *cis*. There is an earlier crystal structure of thiophene-6-azauridine shown in Figure 4-18b that does not include a bromine atom but has not been entered into a crystallographic database.[30] This structure is that of the *trans* conformer. This difference could be due to the presence of the bromine or as the present DFT calculations indicate both conformers exist, it is possible that crystals of either structure can form.

There are also crystal structures of the extended 6-azauridines. The crystal structure of 1a is displayed in Figure 4-19. It is the *trans* conformer that exists in all these structures, consistent with the DFT calculations.

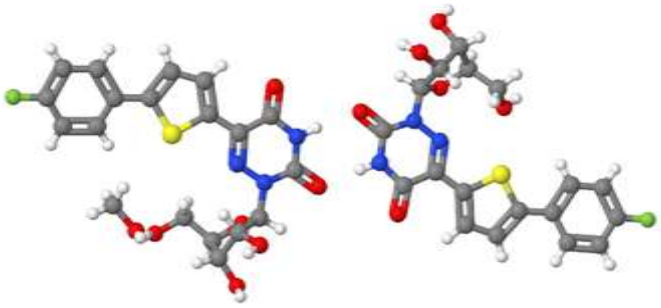


Figure 4-19. Crystal structure of 1a reproduced from Reference 3.

DFT calculations indicated that the substituted phenyl ring adopts a 30° rotation relative to the rest of the molecule when in the ground state. In the crystal structures there is a degree of heterogeneity in the angle of this group. In the crystal structure of 1a the phenyl ring is rotated upwards by 16° in one molecule and by 10° in the other. In the crystal structure of 1f the phenyl ring is rotated upwards by 8° in one molecule and downwards by 12° in another. The difference in energy between a planar at 30° angle structure is minimal, crystal packing could result in distortion of the angle.

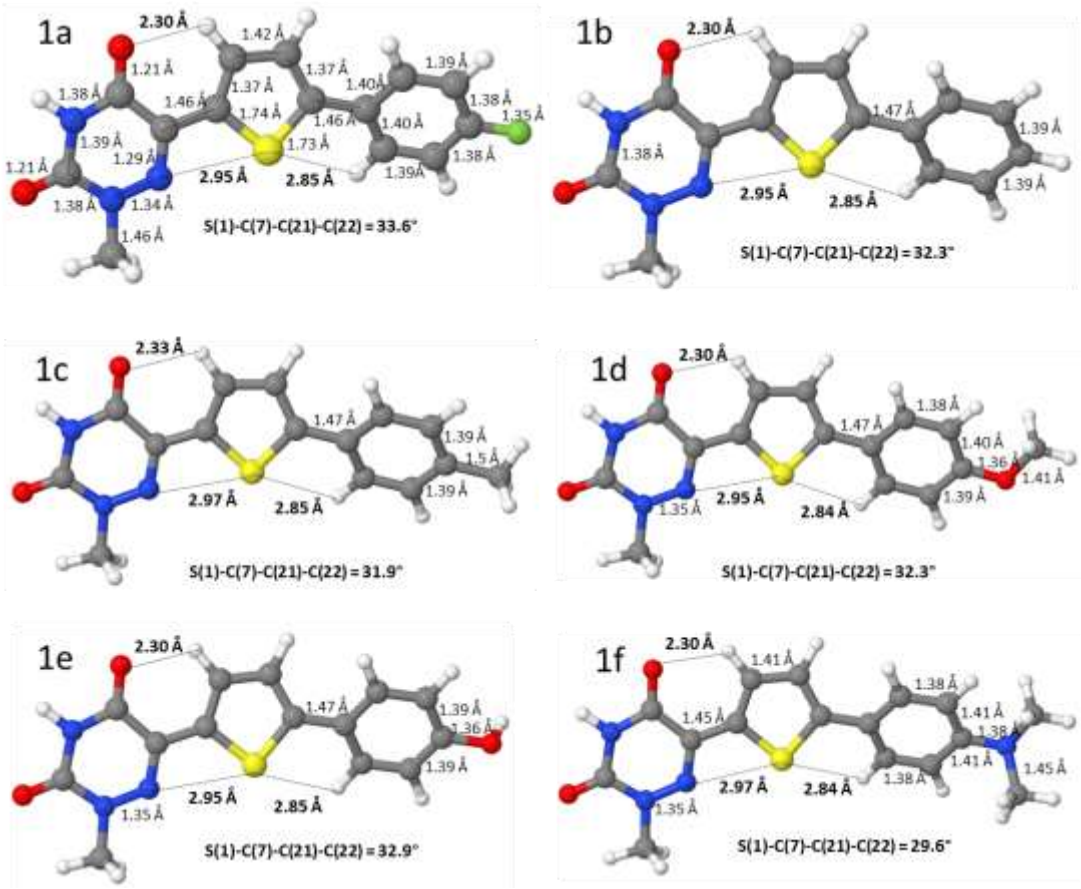


Figure 4-20. Ground state optimised structures of the extended 6-aza-uridines. All bond lengths are illustrated for 1a and then those bond lengths that differ from those in 1a are annotated in subsequent structures. All images created in Jmol.

The ground state optimised structures for the extended 6-azauridines are illustrated in Figure 4-20. 1a has been annotated with: bond lengths; S(1)-C(7)-C(21)-C(22) angle; through space distances between the thiophene sulphur and the 6-azauridine nitrogen; the distance between sulphur and nearest phenyl hydrogen and the 6-azauridine oxygen and thiophene hydrogen. Bond lengths have been included for 1b-1f where they differ from 1a. The greatest difference found between the structures is in the lengths of the bonds neighbouring the Y substituent in the phenyl ring. In 1a these bonds are shortest, 1.38Å whereas in 1f they are longest, 1.41 Å.

Overall the substituent appears to have only a small influence on the overall molecular structure.

4.3.4 Conclusions

The parent molecule, thiophene-6-azauridine, has been shown to exhibit two fluorescence lifetimes in a range of solvents. By comparing average lifetimes to quantum yields it was found that in methanol, ethanol and water, species with lifetimes too short to detect were also present. DFT calculations have shown that in the ground state there are two rotational conformers. It is likely that these isomers are both emissive in dioxane, the *cis* conformer responsible for the 0.1 ns lifetime and the *trans* conformer responsible for the 6.7 ns lifetime. In buffer, methanol and ethanol the *cis* conformer has a lifetime too short to detect, instead the two lifetimes detected can be assigned to the *trans* conformer and a second species, most likely a tautomer of the *trans* conformer. It has been shown that in the excited state that the *trans* and *cis* conformers do not interconvert as the timescale over which this would occur is several orders of magnitude longer than the lifetime of the excited state. More investigation into the tautomer, suspected to be responsible for the second lifetime observed in protic solvents, is required. Calculations into the relative ground and excited state energy of possible tautomers should be carried out. It is also important to continue the excited state calculations related to rotational isomerism by investigating the effect of solvent. In the ground state solvent affected both the relative energy of the two conformers and the height of the barrier between them. It is possible that a similar effect could be seen in the excited state. It is also possible that the effect of solvent on the excited state could be significantly more pronounced. For the molecular rotor DMABN it is known that in protic solvents, in the excited state, the twisted transition state is lower in energy than the planar structure, which results in emission from both the planar and twisted states. In aprotic solvents the twisted state is less stabilised and only emission from the planar conformation is observed.

DFT calculations were also used to investigate the difference between thiophene-6-azauridine and furan-modified 6-aza-uridine. It was found that, unlike thiophene-6-azauridine, furan-modified 6-aza-uridine exists as a single planar *trans* ground state species and likely interconverts to a *cis* planar structure in the excited state. These differences between the furan and thiophene parent molecule explain the different fluorescence properties previously observed in viscous solutions and the lower quantum yield observed for furan-6-azauridine.

Time-resolved measurements revealed that the extended-6-azauridines have two or three component decays, dependent on the nature of the Y-substituent and the solvent they are in. Lifetimes were found to originate from the same rotational isomers as found for the parent

thiophene-6-azauridine. In dioxane, τ_1 was assigned to the *cis* conformer and τ_3 to the *trans* conformer. A tautomer has been suggested as a possible source of the third lifetime, same as observed for the parent. Much like the parent compound it would be instructive to perform calculations that include solvent in order to investigate the effect.

In terms of future use there are several of the extended-6-azauridines that show promise. The most promising could be 1d as it has a high quantum yield in dioxane and a low quantum yield in water. This suggests that it could have significantly different properties when stacked or un-stacked in a DNA duplex. It could therefore be of great use as a probe of base flipping interactions. 1d exhibits two decay components in both dioxane and buffer, this could make lifetime measurements the least complex of all the extended-6-aza-uridines to analyse. As the multiple lifetimes originate from ground state species there is the possibility that incorporation into an oligonucleotide could affect the relative energy of the two conformers, such that one becomes favoured over the other. In order to fully assess its potential use as a probe, measurements of 1d in an oligonucleotide are required.

4.4 References

1 A. Fin, A. R. Rovira, P. A. Hopkins and Y. Tor, in *Modified Nucleic Acids*, eds. K. Nakatani and Y. Tor, Springer International Publishing, Cham, 2016, pp. 1–26.

2 N. J. Greco and Y. Tor, *Tetrahedron*, 2007, **63**, 3515–3527.

3 P. A. Hopkins, R. W. Sinkeldam and Y. Tor, *Org. Lett.*, 2014, **16**, 5290–5293.

4 R. W. Sinkeldam, P. Hopkins and Y. Tor, *Chemphyschem*, 2012, **13**, 3350–6.

5 F. Seela and P. Chittepu, *J. Org. Chem.*, 2007, **72**, 4358–4366.

6 R. W. Sinkeldam, A. J. Wheat, H. Boyaci and Y. Tor, *Chemphyschem*, 2011, **12**, 567–570.

7 M. A. Haidekker, M. Nipper, A. Mustafic, D. Lichlyter, M. Dakanali and E. A. Theodorakis, in *Advanced Fluorescence Reporters in Chemistry and Biology I: Fundamentals and Molecular Design*, ed. A. P. Demchenko, Springer Berlin Heidelberg, Berlin, Heidelberg, 2010, pp. 267–308.

8 D. Rappoport and F. Furche, *J. Am. Chem. Soc.*, 2004, **126**, 1277–1284.

9 M. A. Haidekker and E. A. Theodorakis, *J. Biol. Eng.*, 2010, **4**, 1–14.

10 N. Ghoneim, *Spectrochim. Acta - Part A Mol. Biomol. Spectrosc.*, 2000, **56**, 1003–1010.

11 B. R. Beno, K. S. Yeung, M. D. Bartberger, L. D. Pennington and N. A. Meanwell, *J. Med. Chem.*, 2015, **58**, 4383–4438.

12 D. J. Pascoe, K. B. Ling and S. L. Cockroft, *J. Am. Chem. Soc.*, 2017, **139**, 15160–15167.

13 X. Cao, J. Liu, P. Hong, G. Li and C. Hao, *J. Photochem. Photobiol. A Chem.*, 2017, **346**, 444–451.

14 T. Suhina, S. Amirjalayer, S. Woutersen, D. Bonn and A. M. Brouwer, *Phys. Chem. Chem. Phys.*, 2017, **19**, 19998–20007.

15 F. Zhou, J. Shao, Y. Yang, J. Zhao, H. Guo, X. Li, S. Ji and Z. Zhang, *European J. Org. Chem.*, 2011, 4773–4787.

16 H. Zhu, J. Fan, M. Li, J. Cao, J. Wang and X. Peng, *Chem. - A Eur. J.*, 2014, **20**, 4691–4696.

17 N. Gupta, S. I. Reja, V. Bhalla, M. Gupta, G. Kaur and M. Kumar, *J. Mater. Chem. B*, 2016, **4**, 1968–1977.

18 T. Kobayashi, Y. Harada, T. Suzuki and T. Ichimura, *J. Phys. Chem. A*, 2008, **112**, 13308–13315.

19 T. Kobayashi, H. Kuramochi, Y. Harada, T. Suzuki and T. Ichimura, *J. Phys. Chem. A*, 2009, **113**, 12088–12093.

20 T. Kobayashi, H. Kuramochi, T. Suzuki and T. Ichimura, *Phys. Chem. Chem. Phys.*, 2010, **12**, 5140.

21 M. K. Shukla and J. Leszczynski, *J. Biomol. Struct. Dyn.*, 2007, **25**, 93–118.

22 J. Lorentzon, M. P. Fuelscher and B. O. Roos, *J. Am. Chem. Soc.*, 1995, **117**, 9265–9273.

23 X. Hua, L. Hua and X. Liu, *Phys. Chem. Chem. Phys.*, 2016, **18**, 13904–13911.

24 Gaussian 09, Revision E.01, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox, Gaussian, Inc., Wallingford CT, 2013.

25 Jmol: an open-source Java viewer for chemical structures in 3D. http://www.jmol.org

26 ArgusLab 4.0.1 M. Thompson and Planaria Software, 1997- 2004.

27 B. Mennucci, *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, 2012, **2**, 386–404.

28 T. Yanai, D. P. Tew and N. C. Handy, *Chem. Phys. Lett.*, 2004, **393**, 51–57.

29 C. Adamo and D. Jacquemin, *Chem. Soc. Rev.*, 2013, **42**, 845–856.

30 I. Basnak, M. Sun, T. a. Hamor, F. Focher, a. Verri, S. Spadari, B. Wroblowski, P. Herdewijn and R. T. Walker, *Nucleosides, Nucleotides and Nucleic Acids*, 1998, **17**, 187–206.

Chapter 5 : Photophysical Characterisation of the Quadracyclic Adenine Family of Fluorescent Base Analogues

5.1 Introduction

The focus of this chapter will be a time-resolved fluorescence and computational investigation into the qAN family of fluorescent base analogues. The qAN family consists of four compounds, qAN1 to 4, illustrated in Figure 5-1. These molecules are modified versions of the FBA, quadracyclic adenine (qA).[1] They have been modified by the addition of a nitrogen into the outer six-membered ring. qANs 1-4 differ from each other by the position of the nitrogen in this ring.

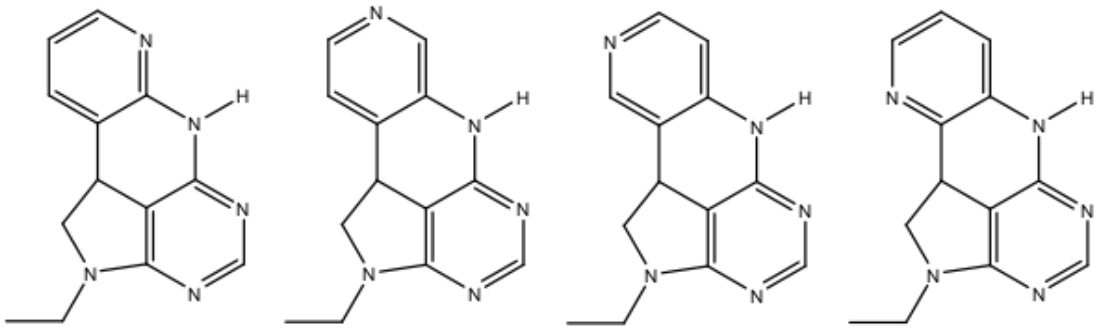


Figure 5-1. Structures of qAN1, qAN2, qAN3 and qAN4 (named left to right).

All four molecules were synthesised by Wilhelmsson *et al.* in 2015.[2] A characterisation of some fluorescence properties was also carried out at this time. This characterisation can be separated into two parts; experimental and computational.

The computational work will be described here first. The qAN family were synthesised based on a TDDFT rationalised design.[3] Oscillator strengths were calculated and used as predictors of quantum yield. Molecules with low $S_0 \rightarrow S_1$ oscillator strengths were predicted to have low quantum yields and molecules with large oscillator strengths, high quantum yields. Synthesis of the qAN compounds seemed to confirm this hypothesis (Table 5-1). Subsequent measurements on 10 other similar molecules showed a similar trend in most cases.[3]

	Calculated λ / nm	Max Experimental λ / nm	Oscillator Strength f	Extinction Coefficient / $M^{-1}cm^{-1}$	Quantum Yield
qAN1	359	360	0.20	9500	0.18
qAN2	360 (337)	340	0.041 (0.22)	9450	0.06
qAN3	359 (324)	330	0.035 (0.18)	5900	0.01
qAN4	355	356	0.30	7300	0.32

Table 5-1. Experimentally and computationally calculated results by Wilhelmsson *et al*. Experimental result in Millipure water. Computational results in PCM water. Value in brackets are the results pertaining to the $S_0 \rightarrow S_2$ transition of qAN2 and qAN3.

Using oscillator strength to predict quantum yield is an unusual method. Oscillator strength is related to extinction coefficient and not normally indicative of quantum yield. Oscillator strengths calculated by DFT can be highly dependent on the choice of functional but are usually reliable for groups of similar molecules.[4,5] By looking at Table 5-1 it can be seen that for the qAN compounds, the oscillator strength of the $S_0 \rightarrow S_1$ transition and extinction coefficient do not correlate. The calculated wavelength of the $S_0 \rightarrow S_1$ transition for qAN1 and qAN4 is almost identical to the experimentally measured wavelength of peak absorbance. The calculated wavelength of the $S_0 \rightarrow S_1$ transition for qAN2 and qAN3 is not identical to the measured peak absorbance wavelength but the $S_0 \rightarrow S_2$ transitions are close to the experimental result. This indicates the qAN1 and qAN4 absorption spectra are dominated by the $S_0 \rightarrow S_1$ transition and that the absorption spectra of qAN2 and 3 are more dominated by the $S_0 \rightarrow S_2$ transitions. The very low oscillator strengths of the $S_0 \rightarrow S_1$ transitions, for qAN2 and qAN3 does explain the low quantum yield of these molecules. A low oscillator strength is indicative of a $n \rightarrow \pi^*$ transition and hence a low quantum yield.

In addition to the absorption spectra and quantum yields already mentioned, further experimental characterisation also took place. Absorption and emission spectra were measured in a range of solvents. A small red shift in emission with increasing polarity was observed for all qAN compounds. The quantum yield of all four compounds was found to be highest in DMSO, but no other quantum yield-solvent trends were observed. Lifetime measurements were made in water, exciting at 377 nm using a diode laser. qAN 1 and qAN 4 were assigned as having single exponential decays. Lifetimes were recorded at two emission wavelengths, 430 nm and 450 nm; these decays were not fitted globally. At 450 nm a ~2 ns component accounting for 7% of the population was also measured for both qAN1 and 4, but this was attributed to scattered light. qAN2 and 3 were found to have two lifetimes, one ~5 ns and one ~ 1 ns.

From steady-state measurements qAN2, and to a lesser extent qAN3, were found to exhibit dual emission in water. When excited at 353 nm, qAN2 was shown to have one minor emission peak at 425 nm and one at 525 nm. At pH 2 only the 525-nm band was present and this peak was attributed to a protonated species. As pH was lowered a change in absorption spectrum was also seen. A second absorption band with a maximum ~420 nm appeared and the 337 nm absorption band slightly decreased in strength and shifted to ~ 350 nm at pH 3.

Although, in water, qAN1 and 4 showed just one emission band, dual emission is seen at more acidic pH. In water, both compounds have emission maxima at around 430 nm; as pH is lowered a second peak appeared at ~500 nm. The protonated species, just as with qAN2, have red-shifted emission and a second red-shifted absorbance band at 420 nm. It was noted that the relative change in peak height with decreasing pH was much more significant than was seen for qAN2. For qAN 1 and 4, at neutral pH, the 430-nm peak dominates, whereas at more acidic pH the 500-nm peak does. For qAN2 it is always the 525-nm peak that dominates, when excited at 353 nm. Appearance of a new absorption band indicates the presence of a new, ground-state species, most likely a protonated species. For qAN1 and 4, as pH is lowered and the amount of this protonated species increases, so too does absorbance at 420 nm and emission at 500 nm. For qAN2 more of the protonated species is formed at lower pH leading to an increase in absorbance at 420 nm, but there is not a significant change in the 430:525 nm peak ratio. It could be that non-protonated qAN2 is virtually non-fluorescent and only emission from protonated qAN2 is seen, a small amount still being present at pH 7. It could also be that the absorbance spectrum of neutral and protonated qAN1 and 4 overlap whereas they are more distinct for qAN2. If excitation of qAN2 was performed at 405 nm, the absorbance peak attributed to the protonated species, as opposed to 353 nm the peak attributed to the non-protonated species, a different emission profile may be found.

Solvent- or pH-dependent dual emission can be indicative of excited-state proton transfer (ESPT). This was suggested as a possible explanation by Wilhelmsson *et al.* for the dual emission exhibited by qAN2 at neutral pH. ESPT generally consists of four basic steps: i) excitation of the ground state species; ii) rapid isomerisation that occurs non-radiatively due to moieties becoming more strongly acidic or basic in the excited state; iii) radiative decay of the new isomer and iv) ground state isomerisation back to the original species.[6] This results in a double-minimum potential-energy surface in both the ground and excited states. The proton transfer involved in excited state isomerisation can occur *via* multiple mechanisms and these can be either intermolecular or intramolecular processes. Kasha, in 1986, was the first to divide these mechanisms into sub-classes.[7] The first sub-class is intrinsic intramolecular proton

transfer. This involves proton transfer across an intermolecular hydrogen bond which is part of a five or six-membered ring. The second sub-class is concerted biprotic transfer which involves proton transfer between molecules in a dimer. The third sub-class is static and dynamic catalysis of proton transfer. Here a second species, which can be solvent, forms a cyclic hydrogen-bonded complex between the hydrogen donating and accepting moieties of the ESPT molecule. Lastly there is proton-relay tautomerisation in which a solvent bridge forms between the hydrogen donor and acceptor. Depending on the height of the energy barrier between the two tautomers, ESPT can occur as proton tunnelling or by the redistribution of vibrational energy. As emission is occurring from a species different to the ground state species being excited, this can result in a large Stokes shift. The relative stability of these species can often be influenced by the environment. Both these features make molecules that undergo ESPT useful molecular probes.

Both intermolecular and intramolecular proton transfer mechanisms have been shown to occur in nucleobases and nucleobase analogues. Often, these are shown to be water-assisted processes. In guanine, for example, keto-enol tautomerism occurs. The mechanism has been shown to be mediated by water. For most guanine tautomers, the transition state of this mechanism involves a guanine cation and a H_3O^+ cation. However, for the N9H tautomer of guanine, the latter is the cation and water acts as an anion.[8] ESPT has also been shown to occur in various base analogues including 1,N^6-ethenoadosine, 8-aza-7-deazaadenine and 8-azapurine[9,10].

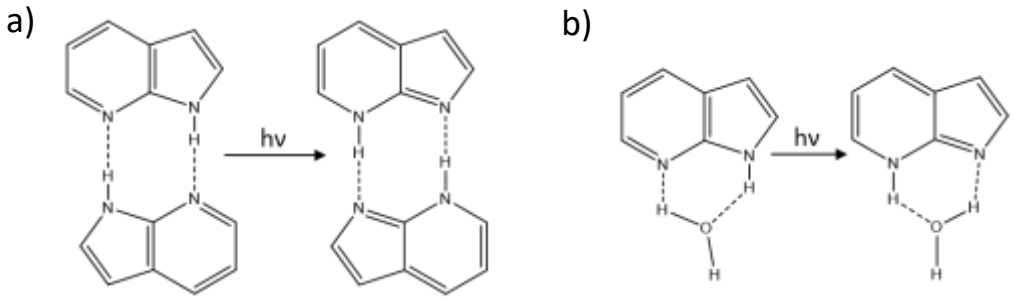


Figure 5-2. Mechanism of excited state proton transfer of 7-azaindole. a) double proton exchange b) proton transfer mediated by water.

For the qAN compounds, excited-state proton transfer would most likely involve the hydrogen of the secondary amine. In this sense, the mechanism is comparable to the well-studied azaindoles. Of the azaindole family, 7-azaindole (7AI) is the most well-studied. It has been used as a fluorescent analogue of tryptophan and a non-base pairing analogue of adenine. It has a quantum yield of 0.03, absorbance at 289 nm and emission at 385 nm in water.[11,12] It undergoes ESPT and has been studied both experimentally and computationally in order to

understand the nature of this. In water, the absorption spectrum is attributed to one tautomer and the emission spectrum to a second. 7-azaindole is known to undergo excited-state proton transfer either by double proton exchange (Figure 5-2a) or by a pathway mediated by solvent molecules. The mechanism by which this occurs with one water molecule is shown in Figure 5-2b. The driving force for this tautomerisation is due to a reversal in relative stability of the two tautomers between the ground and excited states. It is possible that a similar mechanism occurs at the hydrogen bonding face of the qAN molecules, where a similar juxtaposition of nitrogen centres is found.

The tautomerism mechanism proposed for 6-azaindole (6AI) differs from that of 7AI. In this case, the participating nitrogen centres are separated by a greater distance, making both mechanisms illustrated in Figure 5-2 unfeasible.[13] Instead, the proposed mechanism involves an excited-state protonated intermediate. Excited protonation at the N6 nitrogen is followed by excited state deprotonation of the N1. One investigation particularly relevant to work in this chapter, is the study into four azaindole isomers by Cash *et al*, in which 4-azaindole, 5-azaindole, 6-azaindole and 7-azaindole were investigated.[14] They calculated the relative energy of ground and excited state tautomers using TDDFT and found that for 4AI and 5AI, as well as 6 and 7AI, the higher energy ground state tautomer was lower in energy in the excited state.[13,14]

Either ESPT or the presence of multiple ground state species, including protonated species, could account for the dual emission exhibited by the qANs. If the qAN molecules undergo ESPT, emission will be independent of excitation wavelength. Assuming two ground state species do not have identical absorption and emission spectra, excitation at different wavelengths will result in differences in emission spectra. Investigating the emission spectra of the qANs as a function of excitation wavelength could shed light on whether these molecules undergo ESPT, or if the observed dual emission is due solely to the presence of multiple ground state species. Results pertaining to this will be presented in this chapter. Time-resolved measurements can shed light on the number of components present in a sample. Lifetime results for the qAN compounds will also be presented in this chapter. In order to identify if ESPT is likely, and the nature of the tautomers involved, a TDDFT investigation into the qANs has been carried out.

5.2 Materials

qAN 1- 4 bases were received as solids from Prof. L. M. Wilhelmsson, Chalmers University of Technology. Solutions were made using ethanol or phosphate buffer, pH 7.4. Solutions were

made to concentrations of ~ 1 x 10^{-6} M with exact concentrations determined from absorption spectra.

5.3 Results

5.3.1 Steady-state Fluorescence

Excitation and emission spectra of qAN1, qAN2, qAN3 and qAN4 recorded in ethanol are displayed in Figure 5-3. In ethanol, qAN 1 and 4 have absorption maxima at 360 nm and emission maxima at 420nm. Comparatively, the absorption spectra of qAN 2 and 3 are blue-shifted with the absorption maximum of qAN2 at 342 nm, and that of qAN3 at 328 nm. The emission maxima of qAN 2 and 3 are slightly red-shifted compared to qAN 1 and 4, with emission at approximately 430 nm. The emission spectra of all four compounds are independent of the wavelength of excitation. The quantum yield, measured from exciting at the peak absorbance wavelength, in ethanol is slightly higher than in water (Table 5-1); for qAN2, 3 and 4, 0.09, 0.08 and 0.38, respectively, and slightly lower for qAN1, 0.16.

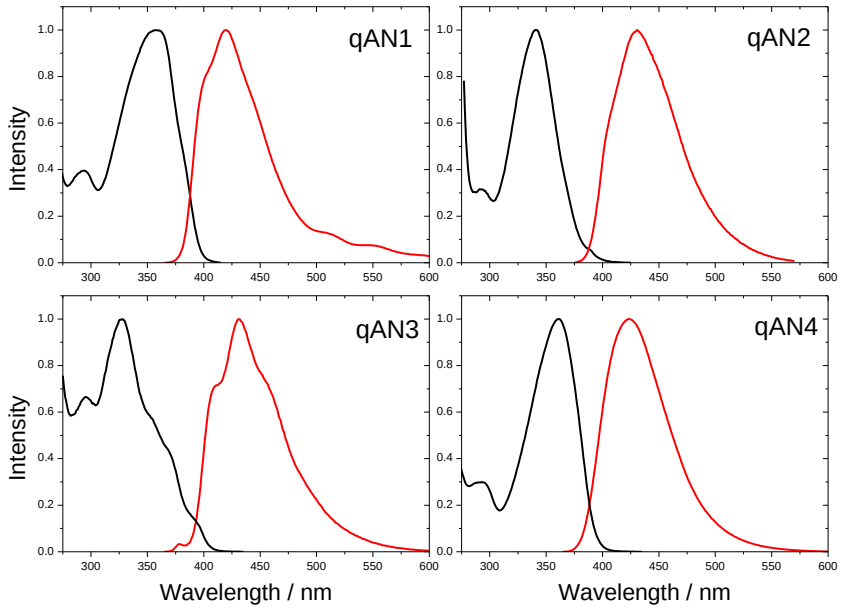


Figure 5-3. Excitation (black) and emission (red) spectra for qAN1-4 in ethanol. Excitation was performed at 340 nm and emission recorded at 430 nm.

Excitation and emission spectra of qAN1, qAN2, qAN3 and qAN4 were also recorded in phosphate buffer and are displayed in Figure 5-4. The resulting spectra and quantum yields are consistent with those reported previously.[3]

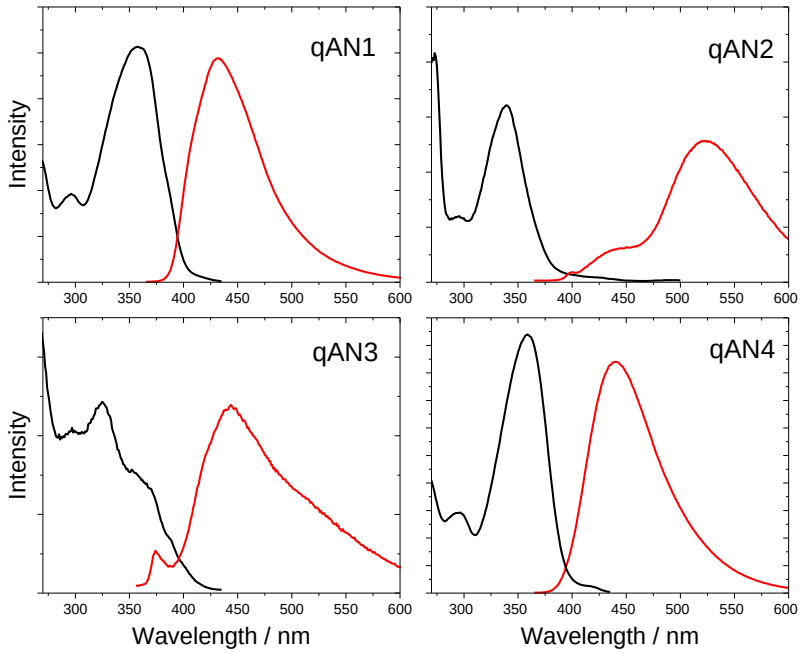


Figure 5-4. Excitation (black) and emission (red) spectra for qAN1-4 in phosphate buffer pH 7.4. Excitation was performed at 350 nm and emission recorded at 450 nm for qAN1, 3 and 4 and at 520nm for qAN2.

The absorption maxima of all four qANs in buffer fall within 1 nm of the corresponding compound in ethanol. The emission spectra in buffer are red-shifted by at least 10 nm compared to ethanol. qAN1 and qAN4 have single emission bands with maxima at ~ 440 nm, qAN3 has a broader spectrum with a peak at 450 nm and qAN2 has dual emission with the major emission peak occurring at 525 nm.

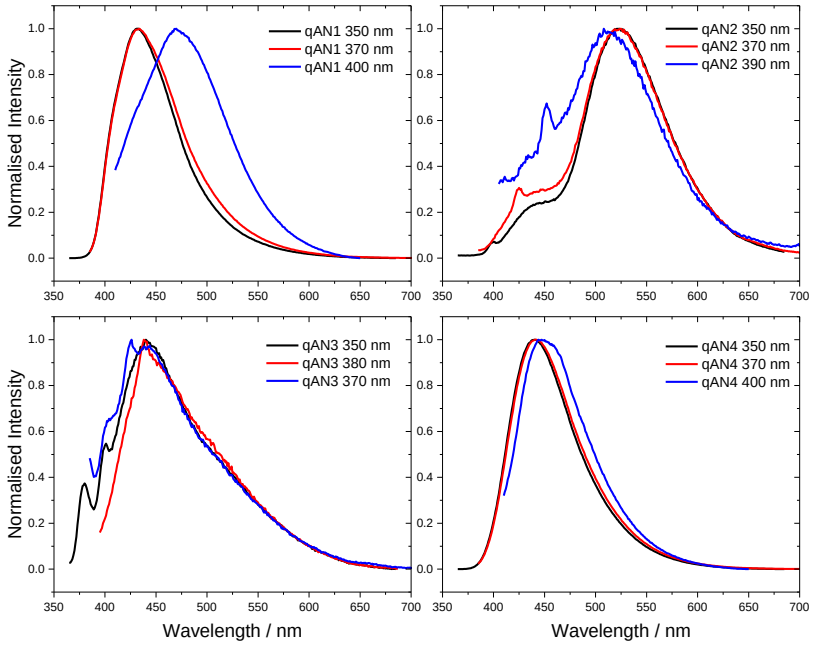


Figure 5-5. Emission spectra of qAN1, 2, 3 and 4 in phosphate buffer pH 7.4 excited at multiple wavelengths.

Unlike in ethanol, emission in phosphate buffer is found to depend on excitation wavelength. For qAN 1 and 4, exciting at the red edge of the absorption spectrum (400 nm) results in red-shifted emission; for qAN1, emission shifts to 470 nm (Figure 5-5), whereas for qAN4 the emission is less red-shifted, to 450 nm. This wavelength-dependence indicates that there are multiple ground-state species present which have distinct emission spectra. For qAN 2 and 3, there is no clear evidence of excitation wavelength-dependence of the emission spectrum. The quantum yield is lower for these compounds resulting in a lower signal-to-noise ratio and more dominant Raman peaks. qAN2 appears to exhibit slightly blue-shifted emission when excited at longer wavelengths and qAN3 appears to exhibits no change (Figure 5-5). This could indicate the presence of a single ground-state species or that all species present have overlapping emission spectra.

5.3.2 Time-resolved Fluorescence

Fluorescence decay parameters for qAN1, qAN2, qAN3 and qAN4 in ethanol (Table 5-2) and phosphate buffer pH 7.4 (Table 5-3) have been recorded. Excitation was performed at 357 nm and decays were collected at three emission wavelengths. A-factors are reported at the peak emission wavelength, results pertaining to the other emission wavelengths are displayed in Appendix III. All four compounds, in both solvents, were found to exhibit multiexponential decay.

	τ_1 / ns	A_1	SS_1	τ_2 / ns	A_2	SS_2	τ_3 / ns	A_3	SS_3	$<\tau>$/ ns
qAN1 (440 nm)	0.15	0.17	0.01	3.1	0.83	0.99	-	-	-	2.65
qAN2 (440 nm)	0.37	0.16	0.01	4.8	0.84	0.99	-	-	-	4.10
qAN3 (440 nm)	0.22	0.20	0.01	3.3	0.10	0.05	8.0	0.70	0.94	5.96
qAN4 (440 nm)	0.70	0.03	0.01	4.6	0.97	0.99	-	-	-	4.50

Table 5-2. Fluorescence decay parameters, fractional contributions to steady-state (*SS*) and average lifetime ($<\tau>$) for qAN1-4 in ethanol excited at 357 nm. A factors (A_i) are shown for peak emission wavelengths (shown in parenthesis in column 1). Global χ^2= 1.06 for qAN1, χ^2= 1.08 for qAN2, χ^2= 1.10 for qAN3 and χ^2= 1.11 for qAN4. Error in τ_1 = 5%, Error in τ_2 = 3%, Error in τ_3 = <1%, Error in A_1 = 5%, Error in A_2 = 3%, Error in A_3 = <1%.

In ethanol, two components are observed in the decays of qAN 1, qAN 2 and qAN 4 indicating the presence of two emitting species (Table 5-2). The major component for qAN 1, 2 and 4 is the longer-lived species, τ_2 which has a lifetime of between 3 and 5 ns. The shorter, sub-nanosecond lifetime component, τ_1, is the minor component, constituting around 20% of the emitting population for qAN1 and qAN2 and only 3% for qAN4. In terms of the contribution to steady-state intensity, this component accounts for 1% or less. Unlike the other three compounds, qAN3 exhibits three components. The major component, τ_3, is a species with an

8 ns lifetime that accounts for 70% of the population. Components τ_1 and τ_2 are similar to those observed for qAN 1.

One component dominates the steady-state spectra for all four qAN compounds when excited at 357 nm. This is consistent with steady-state results which showed no excitation wavelength dependence of emission, indicating that all observed emission was most likely from one species.

In buffer, just as in ethanol, multi-exponential decay was observed (Table 5-3). Unlike in ethanol all compounds show three decay components indicating the existence of three emitting species with distinct decay times.

	τ_1 / ns	A_1	SS_1	τ_2 / ns	A_2	SS_2	τ_3 / ns	A_3	SS_3	$<\tau>$/ ns
qAN1(440 nm)	0.07	0.34	0.01	2.2	0.03	0.02	4.8	0.63	0.97	3.13
qAN2 (540 nm)	0.31	-0.27	0.03	1.5	0.07	0.03	4.6	0.65	0.96	4.30*
qAN3 (450 nm)	0.23	0.21	0.01	4.2	0.14	0.10	8.1	0.64	0.89	5.88
qAN4 (450 nm)	0.09	0.24	0.01	4.5	0.60	0.75	5.3	0.16	0.24	3.53

Table 5-3. Fluorescence decay parameters, fractional contributions to steady-state (*SS*) for qAN1-4 in phosphate buffer (pH7.4). A factors (A_i) are shown for peak emission wavelengths (shown in parenthesis in column 1). $<\tau>$ is the number-average lifetime. * Average of decay components 2 and 3. Global χ^2= 1.09 for qAN1, χ^2= 1.13 for qAN2, χ^2= 1.26 for qAN3 and χ^2= 1.07 for qAN4. Error in τ_1 = 4%, Error in τ_2 = 3%, Error in τ_3 = <1%, Error in A_1 = 4%, Error in A_2 = 3%, Error in A_3 = <1%.

In all cases, except qAN3, the most highly populated species has a lifetime of between 4 and 5 ns. This species dominates the steady-state emission. At the maximum emission wavelength it accounts for ~97% of the steady-state intensity for qAN1 and qAN2 and ~75% for qAN4. The largest decay component for qAN3 has a longer lifetime of 8 ns. This species constitutes >60% of the emitting population and accounts for 88% of the steady-state intensity at the maximum emission wavelength. The 4-5 ns species is the longest lifetime component for qAN1 and qAN2. qAN4 has a second long lifetime component of 5.3 ns which accounts for ~24% of the steady-state emission.

In the case of qAN1 and qAN2, the intermediate lifetime component, τ_2, constitutes only a minor amount, 3% and 7% of the emitting populations, respectively. For qAN1 the population of this component shows a significant dependence on emission wavelength (Appendix III). At 420 nm it accounts for just 1% of the population with this increasing to 9% at 460 nm indicating the presence of a species with red-shifted emission.

Unlike the 6-aza-uridines in Chapter 4, where the source of multiple lifetimes was suggested to be rotational isomerism, the structure of the qAN compounds means it is likely that these multiple lifetimes correspond to different tautomers.

The decay curve of qAN2 differs from that of the other qAN compounds as it shows a rise time, manifesting as a negative A-factor. A rise time indicates that observed emission is not being directly excited. This suggests that the dominant species, the 4.6 ns lifetime component, emitting in this wavelength range is excited indirectly *via* a species with a lifetime of 310 ps. This can be attributed to excited-state proton transfer. The 310 ps component most likely emits in the ~450 nm range, the position at which a small peak is observed in the emission spectrum and where emission dominates in ethanol. The emission spectrum of qAN3 is fairly broad, possibly indicating dual emission. Although at the wavelengths measured no rise time was resolved, the fit of these decays was poor which can be indicative of an unresolved rise time.

The lifetime of the dominant component for qAN2 and qAN4 in ethanol is very similar to the corresponding value in buffer. For qAN1 the lifetime of the dominant component is shorter in ethanol, but still comparable. This could suggest that the same tautomer is predominant in both solvents. Even though the emission spectrum of qAN2 in water is significantly different from the spectrum in ethanol, the decay times of 0.37 ns and 4.8 ns are very similar to the values of τ_1 and τ_3 in buffer. This suggests it is possible that it is the same two species present. For qAN3 the decay parameters in ethanol are very similar to those in buffer. This suggests that the relative populations and emission spectra of the different tautomers emitting in this wavelength region are largely unaffected by the change in solvent.

For qAN2 and qAN3, the quantum yields determined from steady-state measurements are much lower than would be expected from the average lifetimes, in both ethanol and buffer. This indicates there are species present with lifetimes too short to detect at the time-resolution of these measurements. This could indicate that the major ground state species is non-emissive and only a minor ground state tautomer is being detected. This is consistent with calculation performed by Wilhelmsson *et al.* The low oscillator strength of the $S_0 \rightarrow S_1$ transition suggested the presence of a low lying $n \rightarrow \pi^*$ state, indicating the major qAN2 and 3 tautomers would be non-emissive.

5.3.3 Computational Calculations

Fluorescence lifetime measurements have revealed the presence of multiple emitting species, thought to be tautomers. DFT and TD-DFT calculations have been carried out to determine the most likely tautomeric species in both the ground and excited states. The B3LYP functional

and 6-311+g(d,p) basis set were used throughout both the ground-state and excited-state calculations.

5.3.3.1 Ground State

Ground state calculations were performed in order to confirm that the most stable tautomer is the qAN tautomer reported by Wilhelmsson *et al.* and to assess the possibility of other stable ground state species. Calculations were performed using a PCM ethanol solvent. Figure 5-6 illustrates the four different qAN1 tautomers investigated and the nomenclature that will be used to describe them. Similar nomenclature will be used for the other qAN compounds.

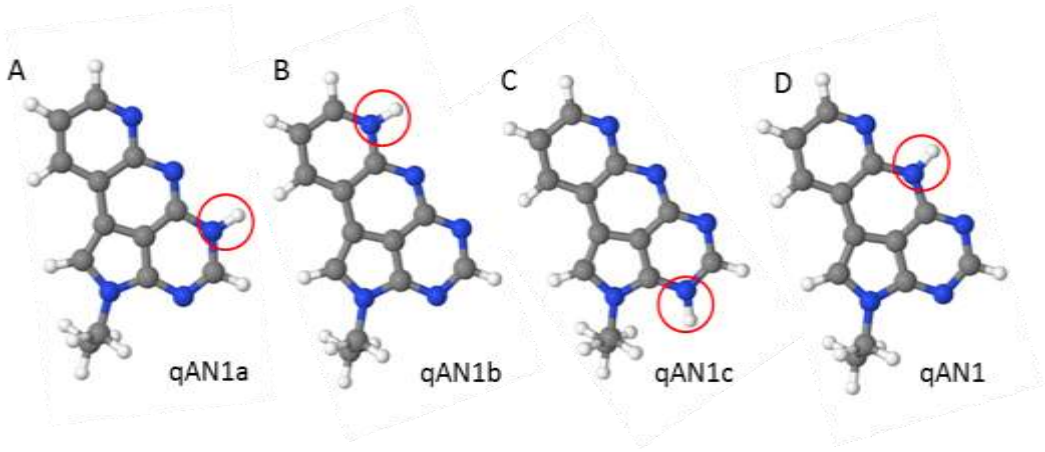


Figure 5-6. Four tautomers of qAN1. A. qAN1a, B. qAN1b, C. qAN1c and D "normal" qAN1. Images created in Jmol.

Each tautomer differs from the others by the nitrogen centre to which the hydrogen is attached. For the qANxb (x denotes 1, 2, 3 or 4) tautomer, the nitrogen on which the hydrogen is attached will be in a different position depending on the member of the qAN family.

Table 5-4 shows the ground-state Gibbs free energy of each of the qANs and their corresponding tautomers. The values are given relative to the Gibbs free energy of the "normal" qAN.

	qANx "normal" Energy/ kJ mol^{-1}	qANxa Energy/ kJ mol^{-1}	qANxb Energy/ kJ mol^{-1}	qANxc Energy/ kJ mol^{-1}
qAN1	0.0	27.9	33.6	61.7
qAN2	0.0	17.0	53.6	49.9
qAN3	0.0	18.6	35.6	51.6
qAN4	0.0	18.1	58.2	50.8

Table 5-4. Gibbs Free Energy / kJ mol^{-1} of each qAN tautomer, relative to the "normal" qAN isomer. x = 1, 2, 3 or 4.

From these calculations it was found that the “normal” qAN tautomer is the most stable species in all cases. All other tautomers are significantly high in energy so as to not exist in the ground state at room temperature. The qANxa tautomer, the tautomer with a hydrogen on the adjacent nitrogen to “normal” qAN, is the second most stable ground state tautomer but was found to exist as < 0.01% of the ground state population.

5.3.3.2 Excited State

The relative energy of these tautomers was also investigated in the excited state. As with the ground state calculation, a PCM ethanol solvent was used. It is known for molecules such as 6-azaindole and 7-azaindole, that the dominant ground state tautomer is not the lowest energy species in the excited state. The absolute Gibbs free energies of the first excited state of the qAN tautomers relative to the excited state energy of the “normal” qAN compound are displayed in Table 5-5. The nomenclature used is the same as that used for the ground state tautomers, here * indicates that this refers to an excited state species.

	qANx* “normal”	qANxa*	qANxb*	qANxc*
qAN1	0.0	127.8	-14.3	9.1
qAN2	0.0	13.4	-39.5	12.1
qAN3	0.0	20.5	-20.5	15.1
qAN4	0.0	-4.4	-27.1	1.3

Table 5-5. Energy of four qAN tautomer (kJ mol^{-1}) relative to the "normal" excited state qAN isomer. x = 1, 2, 3 or 4.

It can be seen from Table 5-5 that, when in the excited state, the “normal” qAN tautomer is no longer the most stable tautomer. In all cases the qANxb* tautomer is the lowest energy. The ground and excited-state energies are represented schematically in Figure 5-7. In this schematic all energies, both ground and excited state, are reported relative to the energy of the normal ground state qAN. The lowest-energy excited-state tautomer is different from the lowest-energy tautomer in the ground state which suggests that for all conformers, excited-state tautomerisation is likely to take place.

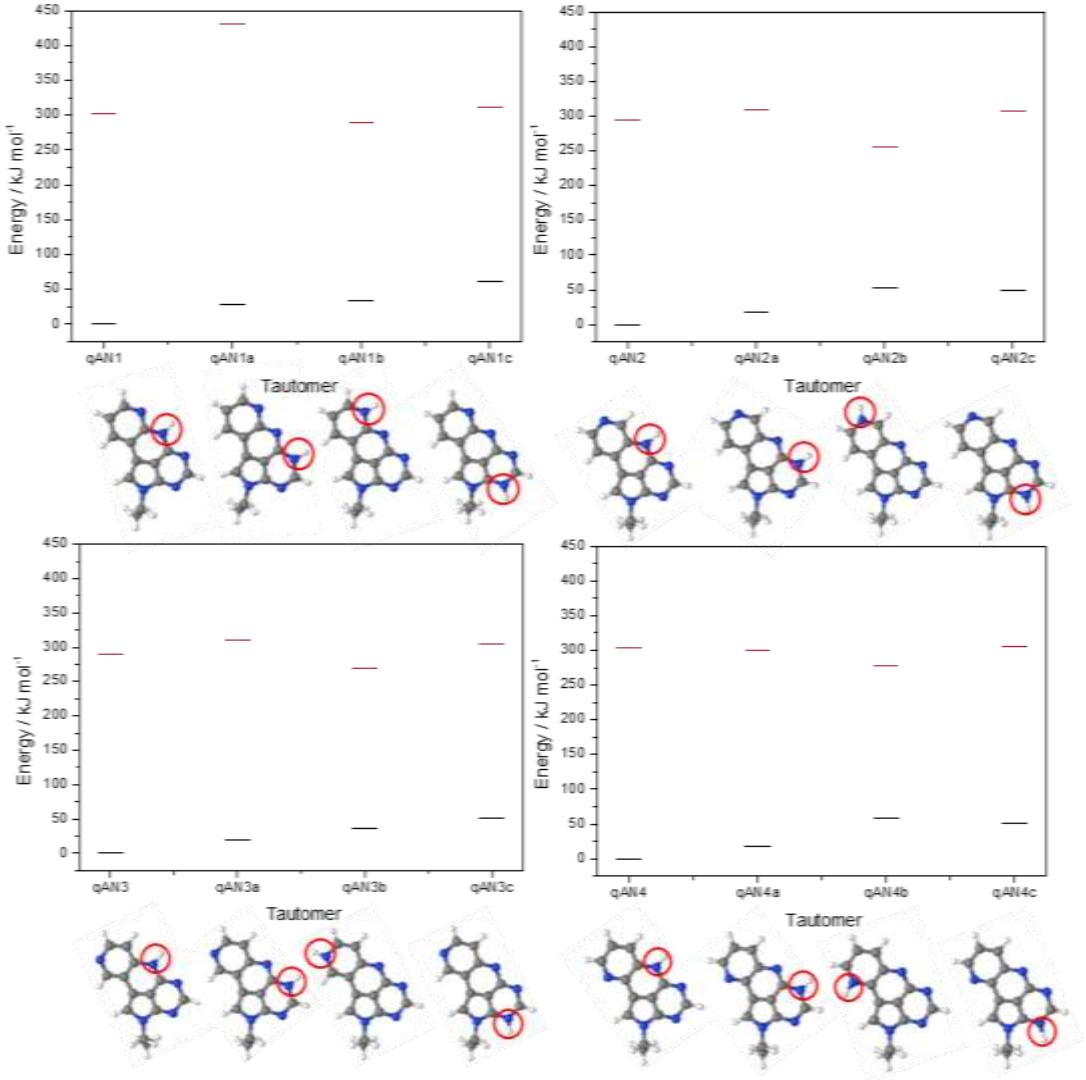


Figure 5-7. Energy of ground (black) and excited state (red) energies of each qAN tautomer relative to energy of the "normal" qAN.

Lifetime measurements for qAN 1, 2 and 4 in ethanol indicate the presence of two emitting species, one that exists to a much greater extent than the other. It is likely one of these species is the "normal" qANx tautomer, the ground state species which is excited and the other is the qANxb* tautomer, formed as a result of ESPT. As there are three lifetimes in the case of qAN3, this cannot be the full story and a third species must be present. It is possible that two tautomers are formed as a result of ESPT. This is not indicated on the basis of the relative energies calculated here but it is possible that inclusion of explicit solvent molecules would have an effect on the relative energies of the tautomers. To fully understand the nature of the mechanism, calculations involving explicit solvent molecules are also required.

For qAN1 the position of the hydrogen in the qAN1 and qAN1b tautomer has the same geometric arrangement as that seen in 7-azaindole. It is therefore possible that a similar mechanism takes place. For qAN2, 3 and 4, the nitrogen groups in question have a different arrangement and a mechanism more similar to the protonated-intermediate mechanism exhibited by 6-azaindole would be required.

5.3.3.3 qAN2 in Water

In water qAN2 exhibits dual emission and a rise time has been resolved from lifetime measurements. This has been attributed to excited state proton transfer. Steady-state measurements at various pH indicate that protonated species have an effect on emission. In acidic solutions the species that emits at 525 nm is present in a greater amount than the species that emits at 420 nm. Calculations have been carried out to help determine where protonation is most likely to occur.

Firstly the relative stabilities of the different qAN2 tautomers in water were investigated (Table 5-6) in the same manner as in ethanol.

	qAN2	qAN2a	qAN2b
Ground	0.0	16.7	50.9
Excited	0.0	9.0	-38.7

Table 5-6. Ground and excited state qAN2 tautomer energies (kJ mol^{-1}) in water.

The stabilities of the qAN2 tautomers in water were found to be similar to the energies seen in ethanol. The most stable tautomer in the ground state is qAN2 and the most stable in the excited state qAN2b*. The ground state energies of qAN2 in water and ethanol are almost identical. In the excited state all isomers are slightly lower energy in water, illustrating the slightly increased stabilisation caused by a more polar solvent. These calculations suggest the same tautomers will be present in both ethanol and water.

It is suggested a protonated qAN2 species exists in water. There are three possible sites at which qAN2 could be protonated. These are indicated in Figure 5-8 and labelled sites A, B and C.

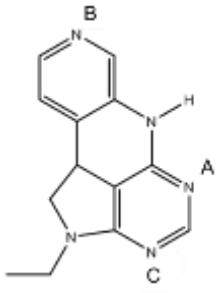


Figure 5-8. qAN2 with possible sites of protonated labelled A, B and C.

Ground state calculations indicate site B is the most favourable position for protonation. Site A is the second most stable and is close enough in energy to site B that both protonated species could exist in a 97:3 ratio. Site B is the same position which has been identified as being the site of the hydrogen in the lowest energy excited state tautomer. It is likely that the additional component seen in lifetime measurements in water is due to this protonated species. It is possible that the protonated species still undergo ESPT to the qAN2b tautomer identified as the most stable excited state tautomer, hence the observation of emission at 525 nm even at low pH. It could be that this process is enhanced at lower pH due to protonation occurring at the same site as that to which the proton is transferred during ESPT.

5.4 Conclusions and Further work

Lifetime measurements indicate multiple emitting species, thought to be tautomers, exist for all members of the qAN family in both water and ethanol. It has also been shown than the decay curve of qAN2 in water shows a rise time. This indicates that the dominant emitting species is being excited indirectly, which can be attributed to excited state proton transfer. Calculations indicate that in ethanol just one qAN tautomer exists in the ground state for each of the qAN compounds. However, the excited state dynamics are more complex. The relative energy of tautomers in the excited state is different to that seen in the ground state, as has been shown to occur for molecules with similar cyclic-amine structures. For all qAN compounds the lowest energy excited state tautomer appears to be tautomer qANxb*.

In summary, in ethanol there is one ground state species, the "normal" qAN tautomer. The second emitting species, τ_2 in qAN 1, 2 and 4, is formed by ESPT. In qAN3 two excited state species are formed by ESPT. In buffer, the same ground state species and tautomer formed by ESPT are observed as in ethanol. The additional lifetime, accounting for less than 10% of the population, observed for qAN 1 and 2 in buffer, can be attributed to either a small amount of a ground-state protonated species or due to a second excited state tautomer. The qAN2 tautomer formed by ESPT is significantly red-shifted compared to the ground state species and rise time is observed when recording at 525 nm. In qAN4, the additional lifetime accounts

for a significant proportion of the emitting population, and is most likely due to a second tautomer formed by ESPT.

The calculations and experiments here are a first step in understanding the nature of the excited state of the qAN compounds. There are a number of experiments that could lead to a deeper understanding. For example, it would be advantageous to measure the fluorescence decay of each compound across more excitation and emission wavelengths in order to potentially resolve rise times in qAN1, 3 and 4. Lifetime measurements made across a range of pH could also help to identify protonated and unprotonated species. This could help identify if these species are present in water.

Further calculations with the inclusion of explicit solvent molecules are also required. Although the calculations performed here indicate excited state isomerisation may occur, the mechanism by which this arises is still unknown. Proton transfer could occur *via* a solvent bridge or could be the result of a protonated intermediate. As the most favourable site for protonation of qAN2 is the same site involved in the lowest energy excited state tautomer, a protonated intermediate is a possibility here.

All characterisation of the qAN compounds has thus far been carried out on the free base. Lifetime measurements in oligonucleotides would also be useful. Use has been made of probes that only exhibit ESPT in certain environments. Should qAN2 exhibit dual emission and a distinct rise time in solution, but not in a duplex, this could be used to distinguish between different DNA environments.

5.5 References

1 A. Dierckx, F.A. Miannay, N. Ben Gaied, S. Preus, M. Björck, T. Brown and L. M. Wilhelmsson, *Chemistry*, 2012, **18**, 5987–97.

2 B. Dumat, M. Bood, M. S. Wranne, C. P. Lawson, A. F. Larsen, S. Preus, J. Streling, H. Gradøn, E. Wellner, M. Grøtli and L. M. Wilhelmsson, *Chem. Eur. J.*, 2015, **21**, 4039–4048.

3 A. F. Larsen, B. Dumat, M. S. Wranne, C. P. Lawson, S. Preus, M. Bood, H. Gradén, L. M. Wilhelmsson and M. Grøtli, *Sci. Rep.*, 2015, **5**.

4 M. Caricato, G. W. Trucks, M. J. Frisch and K. B. Wiberg, *J. Chem. Theory Comput.*, 2011, **7**, 456–466.

5 C. Adamo and D. Jacquemin, *Chem. Soc. Rev.*, 2013, **42**, 845–856.

6 S. Mitra, in *Hydrogen Bonding and Transfer in the Excited State*, 2010, vol. 2, pp. 641–660.

7 M. Kasha, *J. Chem. Soc. Faraday Trans. II*, 1986, **82**, 2379–2392.

8 M. K. Shukla and J. Leszczynski, in *Hydrogen Bonding and Transfer in the Excited State*, eds. K.-L. Han and G.-J. Zhao, John Wiley & Sons, 2011, vol. 2, pp. 125–147.

9 J. Wierzchowski, *Curr. Top. Biophys.*, 2010, **33**, 9–16.

10 J. Wierzchowski, *Nucleosides, Nucleotides and Nucleic Acids*, 2014, **33**, 626–644.

11 Y. Chen, R. L. Rich, F. Gai and J. W. Petrich, *J. Phys. Chem.*, 1993, **97**, 1770–1780.

12 A. V. Smirnov, D. S. English, R. L. Rich, J. Lane, L. Teyton, A. W. Schwabacher, S. Luo, R. W. Thornburg and J. W. Petrich, *J. Phys. Chem. B*, 1997, **101**, 2758–2769.

13 S. M. Twine, L. Murphy, R. S. Phillips, P. Callis, M. T. Cash and A. G. Szabo, *J. Phys. Chem. B*, 2003, **107**, 637–645.

14 M. T. Cash, P. R. Schreiner and R. S. Phillips, *Org. Biomol. Chem.*, 2005, **3**, 3701–3706.

Chapter 6 : Photophysical Characterisation of Pentacyclic Adenine, a Highly Fluorescent Base Analogue

6.1 Introduction

As in previous chapters, Chapter 6 will explore the photophysics of a novel fluorescent base analogue. Unlike Chapters 4 and 5, this chapter will focus solely on one base analogue; pentacyclic adenine (pA). pA is similar to the qAN family of fluorescent base analogues discussed in Chapter 5 in that it is an analogue of adenine and has been designed and synthesised by Wilhelmsson *et al*. pA differs from the qAN compounds by the addition of a six-membered ring and the removal of a heterocyclic nitrogen atom, as can be seen in Figure 6-1.

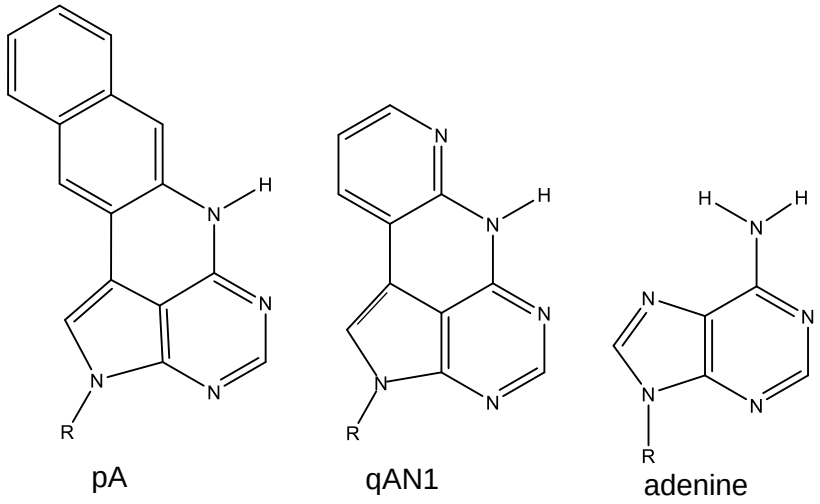


Figure 6-1. Structures of pentacyclic adenine (pA), quadracyclic adenine 1 (qAN1) and adenine.

pA has one of the highest quantum yields of any fluorescent base analogue in water, 0.66. This value is on a par with 2AP and surpasses the brightness of many other FBAs. Unlike 2AP (and most other fluorescent analogues) the quantum yield of pA is high across a range solvents; both polar and non-polar. It is lowest in acetonitrile, 0.64, and highest in dimethyl sulfoxide, 0.84.

Chapters 4 and 5, reported investigations of fluorescent base analogues as either free bases or nucleotides. In this chapter the photophysics of pA, both as a free base and in oligonucleotides, will be explored.

The properties of a fluorescent base analogue in DNA may be different from that of a FBA free in solution, just as this is true for the natural nucleobases themselves. The absorption

coefficients of the natural bases are sensitive to their environment. Absorption coefficients are lowest in double-stranded DNA, slightly higher in single-stranded and highest of all in individual nucleotides.[1] This decrease in absorption coefficient on formation of the duplex is known as hypochromism and is due to interaction of the transition dipole moments of the individual bases in the ordered helical structure. For nucleobases, free in solution, absorption of a photon results in excitation from the ground to a $\pi\pi^*$ state.[2] When in DNA, stacking between bases leads to the formation of so called "collective" excited states.[3] Here the excited state is not localised on one base but delocalised across multiple bases. A charge transfer state (one form of collective excitation) occurs due to neighbouring bases being stacked close enough together such that orbital overlap occurs.[4] Charge (an electron and corresponding hole) are distributed across multiple bases. A further effect is the formation of Frenkel excitons, formed by excited state coupling between the $\pi\pi^*$ states of bases.[5,3] This coupling is dependent on the nature of the base, the degree of base stacking and therefore the form of DNA (A-type or B-type etc.).

For the individual natural nucleobases, the predominant pathway for (the extremely fast) decay from the $\pi\pi^*$ state back to the ground state is internal conversion, due to the presence of a conical intersection between the excited-state and ground state potential energy surfaces.[6,7] For nucleotides in DNA, in addition to deactivation *via* a conical intersection, there are non-radiative transitions between different types of excited states; for example, between a $\pi\pi^*$ and a charge transfer state.[8] The photophysics of the different bases is similar but not identical.

The emission of fluorescent base analogues (as well as some other types of intercalating DNA probes) can be affected by incorporation into DNA.[9] The fluorescence of the well-studied FBA 2AP is quenched between 10 and 100-fold due to base stacking interactions. One mechanism by which excited 2AP is quenched is electron transfer from guanine. Guanine and 2AP must be within approximately 14 Å of each other for this to occur and the nature of the bases between them has an effect on charge-transfer efficiency.[10] The driving force for this electron transfer is the difference in redox potential between guanine and excited 2AP. There is also evidence of charge transfer from the other natural bases to excited 2AP; differences in redox potential between excited 2AP and A, C and T exist, although less than in the case of G.[11] It is therefore accepted that charge transfer is a principal mechanism by which quenching of 2AP fluorescence occurs in oligonucleotides. Excited 2AP acts as an electron acceptor from G or A, whereas hole transfer occurs from C and T to excited 2AP. The mechanism of the quenching of 2AP by the nucleobases is not fully understood and is still an area of current research.

Charge transfer quenching of FBAs by nucleobases is not exclusive to 2AP. An investigation by Narayanan *et al.* into the type of quenching mechanism occurring between the natural bases and the pteridines 6MAP, DMAP, 3-MI and 6-MI was undertaken.[12] Stern-Volmer experiments were carried out to give quenching rates, together with cyclic voltammetry to measure redox potentials. This study elucidated that for 6MAP, DMAP and 3-MI the excited state is most likely quenched *via* electron transfer from surrounding bases; more so from A and G than T and C. 6-MI experiences electron transfer but it is found this occurs only from purines (A and G) with almost no quenching by the pyrimidines. The full quenching mechanism of 6-MI is still under investigation.

When discussing quenching interactions between nucleobases and FBAs it is important to consider the dynamics of DNA and the different stacking conformations a base may adopt. The most well understood base stacking dynamics of any fluorescent base analogue are those of 2AP, as revealed by fluorescence lifetime measurements.[10] When 2AP is in either single or double-stranded DNA, it exhibits a multi-exponential fluorescence decay. Typically four lifetime components are required and each of these represents 2AP in a different stacked environment.[13,14] These lifetimes are summarised in Table 6-1.

	Fluorescence Lifetime (τ_i) /ns				Fractional Amplitude (A_i)			
	τ_1 /ns	τ_2 /ns	τ_3 /ns	τ_4 /ns	A_1	A_2	A_3	A_4
ssDNA	< 0.1	0.2-0.6	2.0-4.0	7.0-9.0	0.50	0.20	0.15	0.15
dsDNA					0.80	0.10	0.05	0.05

Table 6-1. Typical fluorescence lifetimes (τ_i) and corresponding fractional amplitudes (A-factors, A_i) of 2AP in single- and double-stranded DNA oligonucleotides[17].

Each of the lifetimes varies slightly depending on the neighbouring bases, but four distinct lifetimes can be reliably resolved. Lifetimes are similar in both single and double-stranded DNA, with the largest difference between the two being the change in A-factor. Each lifetime can be ascribed to a different stacked conformation and the A-factor to the proportion of this conformation present. The shortest lifetime (τ_1) of 2AP in DNA is attributed to 2AP in the most highly stacked environment. In both ssDNA and dsDNA this well-stacked conformation accounts for most of the emitting population, ~50% in a single strand and ~80% in a double strand. There is a larger proportion of well-stacked 2AP in a duplex, as base pairing reduces the mobility of the bases and favours the population of highly stacked states. The longest lifetime (τ_4) is attributed to extra-helical 2AP. An extra-helical base is the result of base-flipping, a process by which a nucleobase is rotated about the sugar-phosphate backbone so that it lies outside the helix. In this environment 2AP is free from stacking interactions and exists in an environment similar to free 2AP in solution. From the A-factors it can be seen

there is a lower population of extra-helical 2AP in dsDNA compared to ssDNA. Again, this is a reflection of the effect of base-paring in the duplex. The intermediate lifetimes (τ_2 and τ_3) are generally considered to be due to 2AP in partially stacked environments. These conformations are also less populated in the duplex, compared with the single strand.

Although 2AP is the most well studied and understood FBA, the fluorescence decay properties of other fluorescent base analogues have also been studied. One further example will be discussed here, the well-studied guanine analogue, 6-MI. 6-MI does not behave in such a simple manner as 2AP. As a free base it exhibits a bi-exponential decay.[15,16] Unlike 2AP, where the magnitudes of the four lifetimes follow a reproducible pattern (Table 6-1), regardless of sequence, 6-MI lifetimes show more complex sequence-dependence as shown in Table 6-2. A combination of experimental and computational techniques have been utilised to rationalise this more complex behaviour.

	τ_1 /ns	τ_2 /ns	τ_3 /ns	A_1	A_2	A_3
Free 6-MI	-	5.45	6.58	-	0.20	0.80
Purine-rich single strand	0.30	2.21	7.28	0.81	0.14	0.06
Purine-rich double strand	0.21	1.20	5.89	0.83	0.13	0.03
Pyrimidine-rich single strand	-	3.45	7.01	-	0.25	0.75
Pyrimidine-rich double strand	0.26	1.04	6.30	0.88	0.09	0.03

Table 6-2. Lifetimes and A-factors of 6-MI in single and double-stranded DNA, a purine-rich sequence (5 -CACACCTTTTAGA(6-MI)ATCGTCA-3) and a pyrimidine-rich sequence (5 -CACACCTTTTAGAGATC(6-MI)TCA-3). Table modified from reference 18.

The multiple lifetimes exhibited by 6-MI in ssDNA and dsDNA are attributed to different stacked populations; just as with 2AP.[16] The short lifetime components are assigned to well-stacked conformations, the intermediate lifetimes to less well-stacked conformations and the longest lifetime to extra-helical 6-MI. Unlike 2AP the number and magnitude of the lifetimes vary. This is thought to be due to differences in local structure as well as the differences in base stacking interactions.

For 2AP, a reduction in average lifetime (quantum yield) always accompanies a change from ssDNA to dsDNA. In dsDNA a larger proportion of 2AP is in a well-stacked, more highly quenched environment. For 6-MI this is not always the case. In certain sequences an increase in quantum yield is observed on duplex formation. This effect is greatest for the local sequences AT(6-MI)AA and AA(6-MI)TA and has been dubbed “duplex-enhanced fluorescence”.[17] Both the nearest neighbour base (n+1) and the n+2 base affect the enhancement of the fluorescence of 6-MI. Having an adenine at the n+2 position has a significant ability to increase this enhancement. It was hypothesised that this was caused by

either a large extra-helical population of 6-MI in some sequence contexts or restricted dynamic motion. Mukerji *et al.* used a combination of experimental and computational techniques to explain this effect.[18] By measuring the solvent accessible population of a series of DNA strands containing 6-MI, some displaying duplex enhanced fluorescence, a large extra helical population of 6-MI was ruled out. The "duplex enhanced" sequences were in fact found to be slightly less accessible to solvent. MD simulations showed 6-MI strands to be slightly under-twisted compared to the same strands containing guanine. The fluorescence enhanced sequences experienced this under-twisting to a greater extent. The n+2 base effect was explained as it impacts the degree of twisting in the vicinity of 6-MI. Adenine especially is known to form non-B-form structures. Simulations also showed that in all 6-MI sequences, both duplex-enhanced fluorescence and normal, 6-MI took up fewer conformations than guanine. It is evident from 6-MI that the local structure, not just interactions with nearest neighbours, can be important in the photophysics of fluorescent base analogues. It also shows that even for a FBA with a size and structure very similar to a natural base, it may not necessarily be able to access the same conformations as a native base would adopt in a DNA strand.

Electronic energy transfer between FBAs and the natural DNA bases is not a quenching mechanism because, in all cases, the excitation energy of the FBA is lower than that of the natural bases. However, FBAs may be useful as energy acceptors in studying the inter-base energy transfer process, which is believed to play a crucial role in the formation of mutations that can ultimately lead to the development of melanoma. Energy transfer in DNA is therefore currently an area of interest. Due to spectral overlap of the absorption spectra of the natural bases, selective excitation of one base is not possible. The extremely low quantum yield of the natural bases also makes it hard to detect where excitation energy has been transferred to, if at all. To overcome these issues, 2AP has been used as an acceptor of energy transferred from the natural bases. The absorption and emission spectra of 2AP are removed from those of the natural bases allowing selective excitation of the natural bases and selective detection of 2AP. Energy transfer efficiencies from natural bases to 2AP have been investigated in numerous studies. Nordlund, Xu and Evans have been some of the major contributors to this field.[9,19] In a study in 2000, Nordlund and Xu measured a series of 2AP-containing oligos with different sequence contexts to establish efficiencies for transfer from each base.[9] They found significantly more efficient energy transfer from A than any other base. Recently this has been shown to be due, at least in part, to the degree of stacking 2AP undergoes.[20] Energy transfer is more likely in well-stacked conformations and thus energy transfer efficiency depends on both the base and the degree of stacking experienced by the base.

This chapter will explore the steady-state fluorescence properties of, firstly, the pA riboside and, secondly, of a series of pA-containing oligonucleotides. This includes the measurement of energy transfer efficiencies to pA. The latter half of this chapter looks at the time-resolved fluorescence properties of pA and pA-containing oligonucleotides, and discusses the implication of these results on the use of pA as a probe of DNA structure and dynamics.

6.2 Experimental

6.2.1 Materials

The pA base and riboside (pA_r) were received as solids from Prof. L. M. Wilhelmsson, Chalmers University of Technology. Solutions were made to concentrations of ~ 1 x 10^{-6} M with exact concentrations determined from absorption spectra. Measurements were made using the pA riboside unless otherwise stated.

pA-containing 10-mer deoxyoligonucleotides, with sequences as shown in Table 6-3, were received as concentrated stock solutions from the Wilhelmsson group, Chalmers University of Technology. Samples were diluted with phosphate buffer saline (PBS), pH 7.5, 123 mM Na^+. DNA was annealed by combining each modified single strand with 10% excess of its complementary strand at room temperature followed by heating to 95 °C and leaving to cool overnight. In all complementary strands pA was paired with T.

Sequence name	DNA sequence	ϕ_{single}	ϕ_{double}
AT	5'-d(CGCA**A(pA)T**TCG)-3'	0.15	0.17
CC	5'-d(CGCA**C(pA)C**TCG)-3'	0.04	0.04
GA	5'-d(CGCA**G(pA)A**TCG)-3'	0.42	0.14
GG	5'-d(CGCA**G(pA)G**TCG)-3'	0.28	0.13
TT	5'-d(CGCA**T(pA)T**TCG)-3'	0.03	0.11

Table 6-3. Pentacyclic adenine containing deoxyoligonucleotides (10mers). Sequences differ only by the neighbouring bases of pA, and it is these neighbouring bases that give their name to the sequence. Quantum yields in single-strands (ϕ_{single}) and double-strands (ϕ_{double}), excitation 353 nm as calculated by Wilhelmsson *et al* also given.[21]

6.2.2 Energy Transfer Measurements

Absorbances were measured using a Varian Cary 300 UV-visible Spectrophotometer. Emission spectra were measured at excitation wavelengths of 260 nm and 360 nm on the FluoroMax-3P and used to calculate energy transfer efficiencies.

Transfer efficiency, η(t), from a donor, in this case the natural DNA bases, to an acceptor, here pA, following excitation at wavelength λ_{ex} can be calculated using Equation 6-1. This equation was formulated by Nordlund and Xu and used to calculate energy transfer efficiencies from the natural bases to 2AP. Excitation is performed at 260 nm as at this wavelength the natural bases have a maximum, and pA has minimum absorbance.

$$\boldsymbol{\eta(t) = \frac{A_a(\lambda_{ex})}{A_d(\lambda_{ex})} \times \left(\left(\frac{F(\lambda_{ex}) Q_{pA}}{F_a(\lambda_{ex}) Q_{pAoligo}} \right) - 1 \right)}$$

Equation 6-1

where $A_a(\lambda_{ex})$ is the absorbance of the energy acceptor pA at the excitation wavelength $\lambda_{ex,}$ here 260 nm. $A_d(\lambda_{ex})$ is the absorbance of the energy donors, the natural bases, also at 260nm. $A_d(\lambda_{ex})$ is calculated by subtracting the absorbance of pA riboside, at an identical concentration, from that of the oligonucleotide sample. $F_a(\lambda_{ex})$ is the integrated fluorescence emission intensity of pA when directly excited at 260 nm and $F(\lambda_{ex})$ is the integrated emission intensity of the pA-containing oligonucleotide when excited at 260 nm. Q_{pA} is the quantum yield of pA and Q_{pAoligo} is that of pA in the oligonucleotides. The calculated efficiency is the total for all the bases. To calculate the average efficiency per base this must be divided by the number of bases in the oligonucleotide.

6.3 Results and Discussion

6.3.1 Steady-state Spectroscopy

6.3.1.1 pA Riboside

The absorption, excitation and emission spectra of pA_r in two alternative solvents, phosphate buffer saline (PBS) and ethanol, are shown in Figure 6-2.

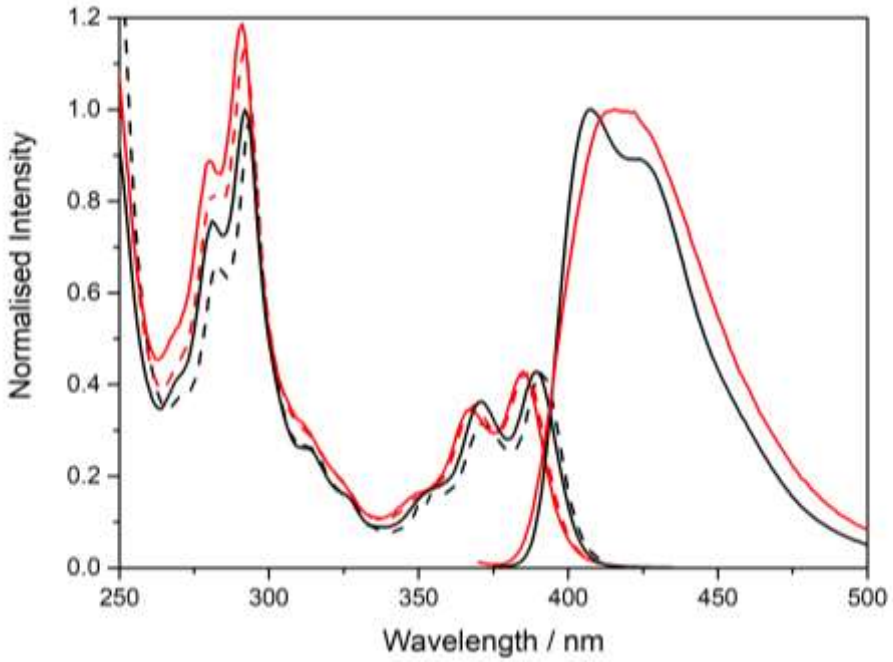


Figure 6-2. Absorption (dotted line), excitation (solid line) and emission spectra of pA riboside in PBS (red) and ethanol (black). Excitation performed at 360 nm and emission collected at 450 nm.

There are notable differences between the spectra in the two solvents. Firstly, there is a difference in emission spectral profile. While the emission spectrum in PBS is unstructured, in ethanol it has distinct features, a peak at 405 nm and a shoulder at 425 nm. Secondly, there is a difference in the wavelengths of both excitation and emission maxima. There is a larger Stokes shift in PBS, resulting from a blue shift in the excitation spectrum and a red shift in the emission spectrum, relative to the respective spectra in ethanol. The larger Stokes shift in aqueous solvent can be attributed to solvent relaxation and implies that the excited state of pA is more polar than the ground state.

The excitation spectrum of pA in PBS is blue-shifted relative to that of pA in the less polar solvent, ethanol. This signifies a destabilisation of the excited state in PBS relative to the ground state. This occurs because the solvent is static during excitation to the Frank-Condon state and the ground state arrangement of PBS is less favourable than that of ethanol. A red shift of the emission spectrum of pA in PBS can again be attributed to the relative stabilising effects of each solvent. Emission occurs from the relaxed excited state, after reorganisation of solvent molecules has occurred. Therefore, as PBS is more polar, it has a greater stabilising effect on the polar excited state of pA than ethanol. Consequently, the relaxed excited state in PBS is lower.

Despite the differences in the spectral profile, the quantum yield of pA in water and ethanol is reported to be almost identical, 0.662 in water and 0.664 in ethanol.[21]

Comparison of the absorption and excitation spectra displayed in Figure 6-2 show small differences. Both spectra have two peaks, a dominant peak at 390 nm and a second peak at 370 nm. The intensity ratio of these peaks differs between absorption and excitation spectra. This is observed to a greater extent in ethanol than PBS. In each solvent there is also a small shift in the absorption spectrum to a longer wavelength, compared with the excitation spectrum. These observations suggest that there are species present that are not contributing to the emission at the wavelength (450 nm) at which the excitation spectra were recorded. The existence of more than one species is supported by closer examination of the excitation spectrum which reveals a dependence of the intensity ratio of the two excitation bands on the emission wavelength, as shown in Figure 6-3.

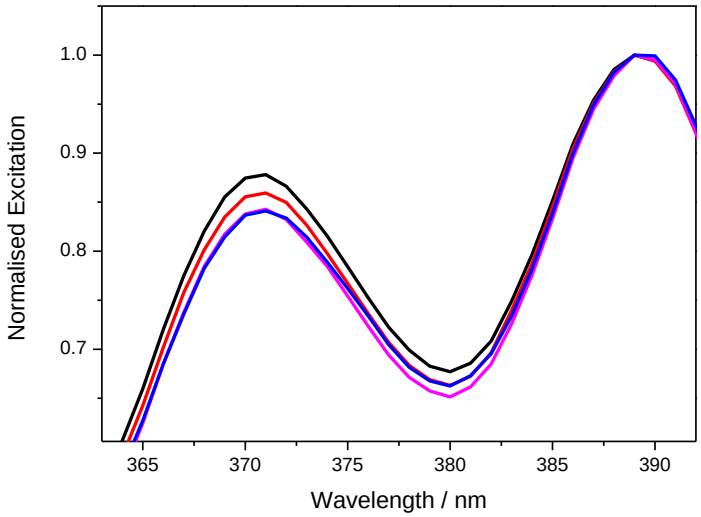


Figure 6-3. Excitation spectra recorded at an emission wavelength of 420 nm (black), 440 nm (red), 460 nm (pink) and 480 nm (blue).

6.3.1.2 pA-containing Oligonucleotides

The absorption spectra of free pA riboside and pA incorporated into oligonucleotides are displayed in Figure 6-4. The absorption maximum of pA incorporated into oligonucleotides, is red-shifted compared to that of pA riboside in PBS. The oligonucleotide spectra are more similar to that of pA in ethanol.

The red-shift observed on incorporation into an oligonucleotide is most likely due to the polarity difference between buffer and the local environment in the oligo. In the oligonucleotides, when flanked by different bases, a slight variation in the absorption maximum of pA is seen. In oligo AT, pA exhibits the shortest wavelength absorption with a

maximum at 389 nm; in TT, GA and GG, pA has an absorption maximum at 390 nm. In CC, pA has a maximum absorbance at 391 nm, the same as that observed for pA in ethanol. These sequence-dependent effects suggest that the local polarity of the environment experienced by pA in the oligonucleotides depends on the identity of the neighbouring bases.

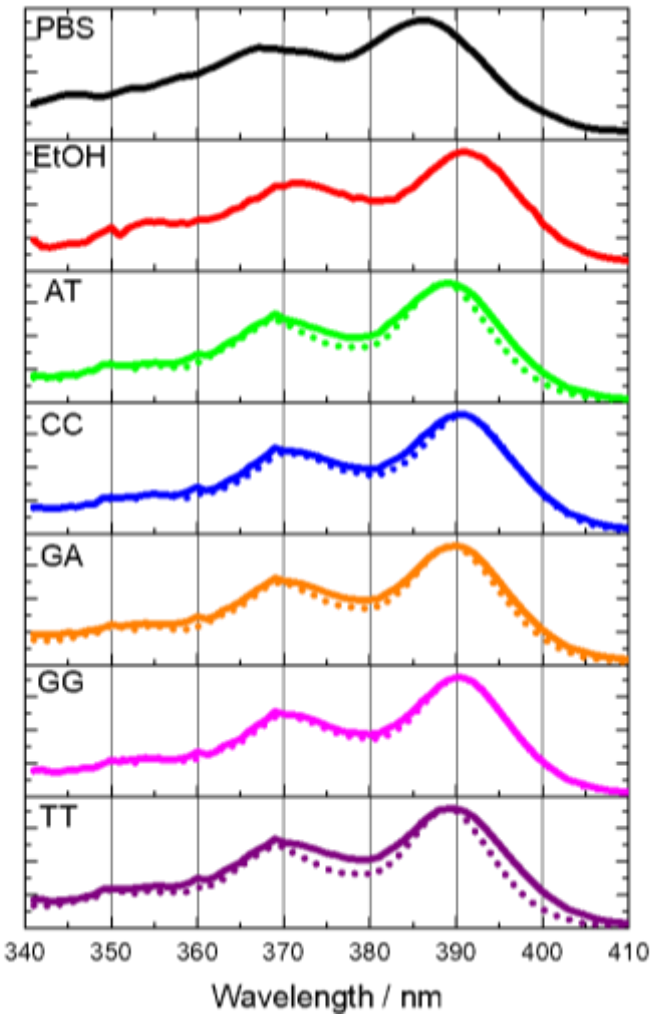


Figure 6-4. Absorption spectra of pA riboside in PBS (black), pA riboside in ethanol (red) and pA containing oligonucleotides with different combinations of nearest neighbours. AT (green), CC (blue), GA (orange), GG (pink) and TT (purple). Solid line denotes single-stranded and dotted line denoted double-stranded oligonucleotides.

Figure 6-4 also illustrates the difference in pA absorption spectra between single and double-stranded oligonucleotides. In oligos CC GA and GG, the absorption spectra of single and double strands are essentially identical. For AT and, TT, a slight difference is observed between single and double strands; in the duplex, the two absorption peaks are narrower.

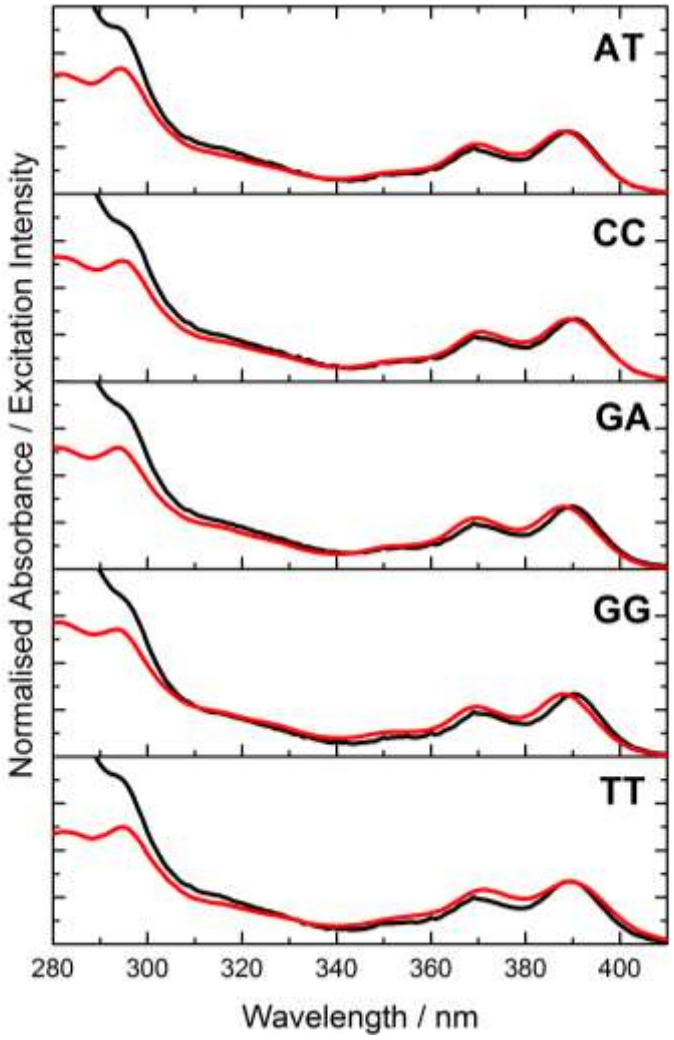


Figure 6-5. Normalised excitation spectra (red) and absorption spectra (black) of pA in single-stranded oligonucleotides with different combinations of nearest neighbours. Excitation recorded at an emission wavelength of 420 nm.

Just as for the pA riboside, in the range 410-340 nm excitation spectra are similar but not identical to the absorption spectra recorded for pA in the five oligonucleotide sequences. This is illustrated in Figure 6-5. The excitation maximum at around 390 nm differs by ± 1 nm from the maximum seen in the absorption spectrum but the most noticeable difference is seen in the relative height of the 370 nm and 390 nm peaks. The 370 nm absorption peak is relatively lower than the 370 nm excitation peak, indicating that a species with absorption at 360 nm is contributing to the emitting population more so than the species at 390 nm. At a wavelengths lower than 340 nm large discrepancy between excitation and absorption spectra is observed due to absorption from the natural bases.

The excitation and emission spectra of pA in the five oligonucleotides, both single and double-stranded, are displayed in Figure 6-6.

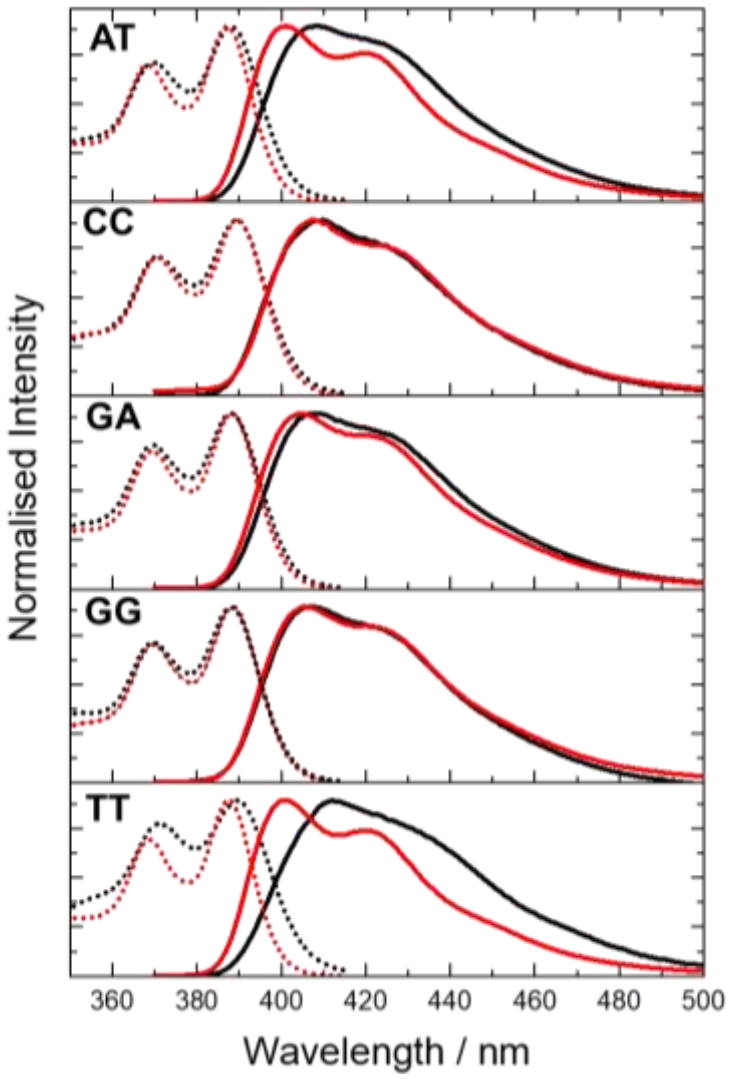


Figure 6-6. Normalised excitation (dotted lines) and emission spectra (solid lines) of pA in single (black) and double-stranded (red) oligonucleotides with different combinations of nearest neighbours. Excitation at 360 nm and emission wavelengths at 420 nm.

Emission spectra for pA in all five of the single-stranded oligonucleotides are similar, although there are some small differences in the intensity of the shoulder at 425 nm. When comparing the emission spectra of pA in double-stranded oligonucleotides, stark differences are observed. This is most noticeable for AT and TT, where the emission is blue-shifted and the spectra are significantly more structured than for the other sequences. There are also differences in the single stranded and double stranded excitation spectra; again this is most noticeable in AT and TT. The difference is observed as a slight difference in peak position and narrowing of the peaks in the double strands compared to the single strands. Notably in TT, but also to a lesser extent in AT and GA, the 370 nm peak is lower relative to the 390 nm peak in the double strand than the single. This indicates that there are differences in emitting populations between the single and double strands.

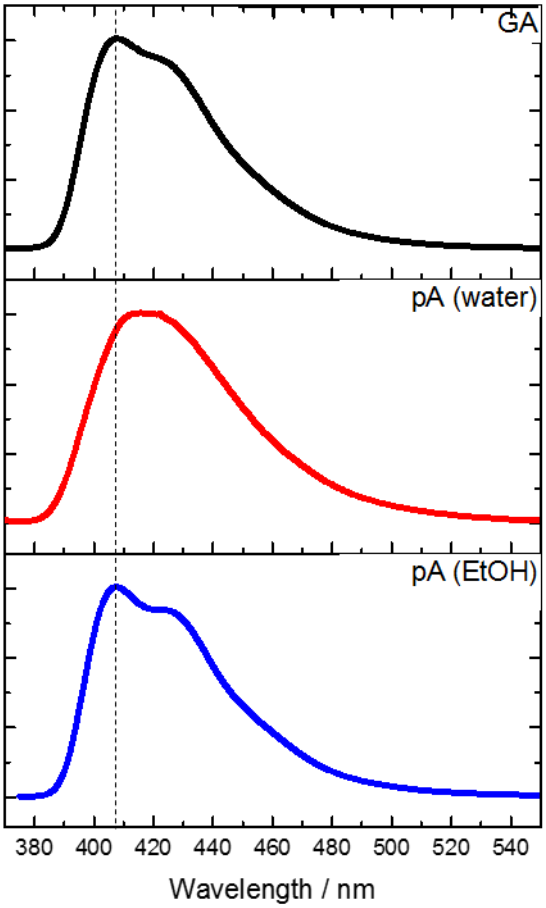


Figure 6-7. Normalised emission spectra of pA in single-stranded oligonucleotide GA (black), in PBS (red) and in ethanol (blue). Excitation wavelength 360 nm.

Figure 6-7 exemplifies the difference in emission spectra between pA riboside in solution and pA in an oligonucleotide.

The structured emission spectra of pA in oligonucleotides more closely resembles the emission spectra of the pA riboside in ethanol than that of pA riboside in buffer. A more structured emission spectrum in oligonucleotides than when free in solution has also been observed for FBAs in the tC family.[22] For tC this effect was attributed to the increased rigidity of the DNA environment compared to tC free in solution, resulting in an increased resolution of vibronic structure, akin to the effect seen when freezing. However, it was not reported whether a similar effect was observed for tC in ethanol. For pA in water, when the viscosity was increased by addition of glycerol, a spectral shape similar to that seen in ethanol or in an oligonucleotide was observed (Figure 6-8). Conversely, increasing the viscosity using sucrose resulted in no such change, even at viscosities much greater than those obtained using glycerol. This suggests the change in spectral profile arose from the addition of alcohol, rather than the increase in viscosity.

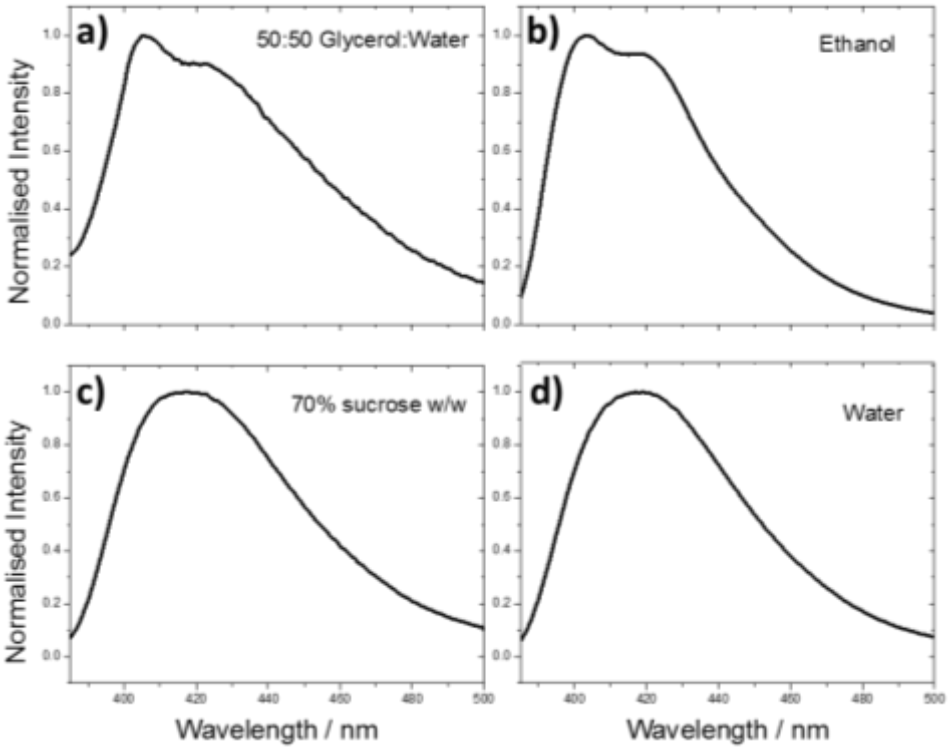


Figure 6-8. a) pA riboside in 50:50 glycerol: water (viscosity = 6 cP, 20 C°).[23] b) pA riboside in ethanol. c) pA riboside in 70% by weight sucrose in water (viscosity = 450 cP, 20 C°).[24] d) pA riboside in water.

If the structure observed in the absorption and emission spectrum were due to the same vibronic progression, the wavenumber difference between the bands should be similar for both spectra. In the absorption spectra difference between the two peaks is 1379 cm^{-1} whereas in the emission spectrum it is 925 cm^{-1}. This discrepancy is too large to be consistent with vibronic structure.

6.3.1.3 Energy Transfer in Single-strand Oligonucleotides

When comparing the excitation spectrum of pA in single-strand oligonucleotides with that of pA_r, a significant difference in the spectra is seen between 255 nm and 280 nm, the wavelength range in which the natural bases have maximal absorbance, as shown in Figure 6-9. The excitation spectrum of pA in an oligonucleotide has significantly higher intensity than that of the pA_r in this region and shows absorption features characteristic of the natural bases; this indicates that there is energy transfer from the natural bases to pA.

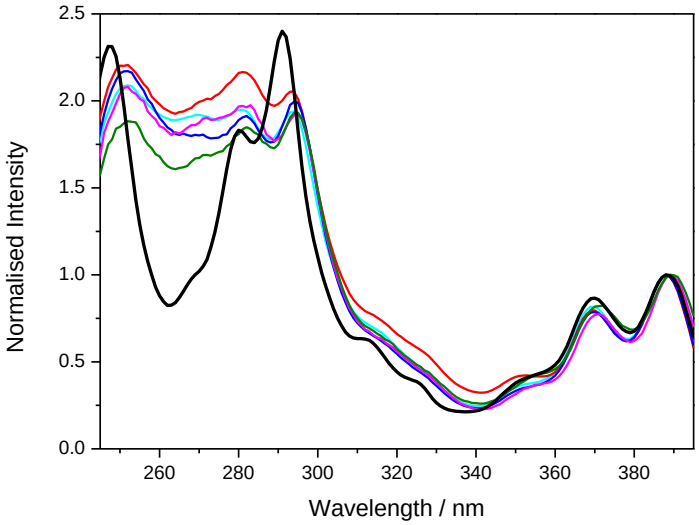


Figure 6-9. Excitation spectra of pA in ethanol (black) and single-stranded oligonucleotides with different combinations of nearest neighbours. GA (cyan), GG (red), AT (blue), CC (pink) and TT (green). Spectra are normalised at the intensity of the ~ 390nm peak. Spectra recorded at an emission wavelength of 420 nm.

Whether exciting pA directly at 360 nm or exciting the natural bases at 260 nm the resulting emission spectrum are almost identical, as shown in Figure 6-10. This indicates that the emitting species is the same in both cases and confirms that excitation of the natural bases at 260 nm results in indirect excitation of pA.

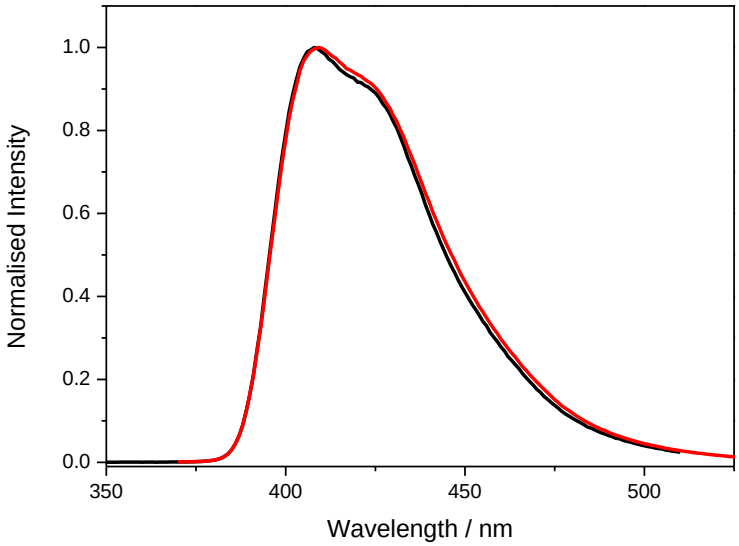


Figure 6-10. Normalised emission spectra of GA excited at 360 nm (red) and at 260 nm (black).

At 260 nm the absorbance of free pA is not zero, but accounts for around 25% of the absorption of the pA oligos at this wavelength. To investigate the extent of indirect excitation, the average energy transfer efficiency per natural base was determined for each of the single-strand oligonucleotides, using Equation 6-1, as displayed in Table 6-4.

Sequence Name	Energy transfer efficiency (η(t)) % per base
AT	11.9
CC	16.0
GA	12.3
GG	9.7
TT	11.8

Table 6-4. Average energy transfer efficiency per natural base, η(t), for single-stranded oligonucleotides with different combinations of pA nearest neighbours. There is an estimated ± 5% error.

The energy transfer efficiencies per base vary depending on the sequence context of pA. Transfer efficiencies decrease in the order CC > GA≈ AT≈ TT> GG. These results may be compared with previous measurements on 2AP-containing oligonucleotides.[9] In a single-stranded adenine decamer the average transfer efficiency to 2AP was found to be 35% at 25 °C.[9] For similar strands consisting of guanine, cytosine and thymine this value was much lower, 5.1, 5.3 and 3.5 % respectively. These values were measured at 4°C. At room temperature they were found to be over 10% lower. The average efficiency of energy transfer from the natural bases to pA is lower than that for transfer from A to 2AP, but higher than for transfer to 2AP by the other bases. Differences in energy transfer efficiency will not be due solely to the identity of the neighbouring base but can also result from differences in the local stacking orientation. The results shown in Table 6-4 could indicate that pA is in a more unstacked conformation when neighboured by guanine, resulting in a significantly lower energy transfer efficiency.

6.3.2 Time-resolved Spectroscopy

6.3.2.1 pA Riboside

Time-resolved fluorescence decay parameters for the pA riboside in PBS and ethanol are displayed in Table 6-5. The excitation wavelength was 390 nm, corresponding to the longest wavelength band in the excitation spectrum. All emission spectra displayed in this chapter have been collected exciting at 360 nm in order that the full spectrum could be shown. When excitation is performed at 390 nm, the spectrum looks almost identical to the 360 nm.

Solvent	τ_1 / ns	τ_2 / ns	A_1	A_2	SS_1	SS_2	$<\tau>$ / ns
PBS	0.19	5.97	0.35	0.65	0.02	0.98	3.95
Ethanol	0.16	4.82	0.20	0.80	0.01	0.99	3.84

Table 6-5. Fluorescence decay parameters, fractional contributions, steady-state intensity (SS) and average lifetime ($<\tau>$) for pA in PBS and ethanol excited at 390 nm. A-factors (A_i) are shown for an emission wavelength of 430 nm. Global χ^2= 1.07 for pA in PBS and χ^2= 1.09 in ethanol. Error in τ_1 = 5%, Error in τ_2 = < 1%, Error in A_1 = 4%, Error in A_2 = < 1%.

In both solvents, pA exhibits bi-exponential decay. The dominant component has a lifetime of 5.97 ns in PBS and 4.82 ns in ethanol. This component accounts for the majority of the emitting

population 65% and 79%, in PBS and ethanol, respectively. A short lifetime: 0.19 ns in PBS and 0.16 ns in ethanol, accounts for the remaining population. Despite differences in lifetime and relative population, the average lifetimes in PBS and ethanol are similar, 3.95 ns and 3.84 ns. This is consistent with the similar quantum yield values, 0.662 and 0.664 reported previously.

The short lifetime component constitutes less than 2% of the steady-state emission intensity, when excited at this wavelength, i.e. the steady-state spectrum is due almost entirely to the longer lifetime species. This is consistent with the observed discrepancy between the excitation and absorption spectra (Figure 6-2). Comparison of the absorption and excitation spectra suggest that the short-lifetime (low quantum yield) species has an absorption spectrum that is slightly red-shifted compared with that of the long-lifetime (high quantum yield) species.

Decay parameters collected at several emission wavelengths for pA in ethanol are shown in Table 6-6. It can be seen that the A factor of the longer lifetime component slightly increases with increasing emission wavelength. This could suggest that the longer lifetime component has a slightly red-shifted emission spectrum compared to that of the short lifetime component but as the difference is small and a difference is only observed at one wavelength this more likely to be due to error in fitting.

λ_{em} / nm	A_1	A_2	$<\tau>$ / ns
430	0.20	0.80	3.84
450	0.18	0.82	3.99
470	0.18	0.82	3.99

Table 6-6. Fluorescence decay parameters for pA in ethanol, excitation at 390 nm. A factors at emission wavelengths of 430 nm, 450 nm and 470 nm. Globally fitted lifetimes τ_1 = 0.16 and τ_2= 4.82 ns.

The origin of the two species revealed through time-resolved measurements could be due to the presence of two tautomers of pA, with structures as illustrated in Figure 6-11. Similar tautomers have been suggested for tC, a fluorescent base analogue with a similar structure to pA.

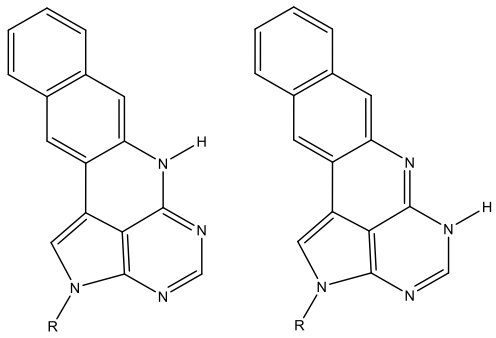


Figure 6-11 A) Structure of pA B) Structure of possible pA isomer.

6.3.2.2 pA-containing Oligonucleotides

The time-resolved fluorescence decay parameters of pA in single-stranded oligonucleotides are presented in Table 6-7.

		τ_1 / ns	τ_2 / ns	τ_3 / ns	A_1	A_2	A_3	SS_1	SS_2	SS_3	$<\tau>$ /ns
GA	410	0.89	-	7.24	0.50	-	0.50	0.11	-	0.89	4.06
	430				0.50	-	0.50	0.11	-	0.89	4.09
	450				0.48	-	0.52	0.10	-	0.90	4.18
GG	410	0.55	2.14	7.48	0.56	0.15	0.29	0.11	0.11	0.78	2.80
	430				0.56	0.15	0.29	0.11	0.11	0.78	2.80
	450				0.54	0.16	0.30	0.10	0.12	0.78	2.89
AT	410	0.16	1.61	6.72	0.56	0.25	0.19	0.05	0.23	0.72	1.76
	430				0.56	0.24	0.20	0.05	0.21	0.73	1.81
	450				0.57	0.24	0.19	0.05	0.22	0.73	1.79
CC	410	0.20	0.78	3.57	0.71	0.27	0.02	0.33	0.50	0.17	0.43
	430				0.69	0.29	0.02	0.31	0.51	0.16	0.45
	450				0.69	0.29	0.02	0.29	0.48	0.23	0.46
TT	410	0.10	0.80	3.96	0.82	0.13	0.05	0.21	0.27	0.51	0.36
	430				0.82	0.13	0.05	0.21	0.27	0.51	0.38
	450				0.82	0.13	0.05	0.21	0.27	0.52	0.38

Table 6-7. Fluorescence decay parameters, fractional contributions to steady-state intensity (SS) and average lifetime ($<\tau>$) for pA in single-stranded oligonucleotides with different combinations of nearest neighbours excited at 390 nm. A-factors (A_i) are shown for emission wavelengths of 410 nm, 430 nm and 450 nm. Global χ^2= 1.20 for GA, χ^2= 1.06 for GG, χ^2= 1.15 for AT, χ^2= 1.20 for CC and χ^2= 1.32 for TT. Error in τ_1 = 10%, Error in τ_2 = 3 % Error in τ_3 = <1 %. Error in A_1 = 2%, Error in A_2 = 1 % Error in A_3 = <1 %.

In the single-stranded oligonucleotide GA, two fluorescence lifetimes are required to adequately describe the decay; a short lifetime component of 0.89 ns accounts for 50% of the population and a longer 7.24 ns lifetime component accounts for the remaining 50%.

In single-stranded GG, AT, CC and TT, three components are required to describe the fluorescence decay and the lifetimes and relative population of each lifetime are highly sequence-dependent. For all four of these sequences, the shortest lifetime component, ranging from 0.55 ns in GG to 0.10 ns in TT, constitutes the largest proportion of the population. In GG, this population accounts for just 56% of the total, but as much as 82% of the population in TT.

A large variation in the long lifetime component is also observed ranging from 6-7 ns in GG and AT to around 3.5-4 ns in CC and TT. The relative proportion of this long lifetime also varies, contributing to a significant proportion of the population in GG and AT but just a few percent of the total population of CC and TT. Looking now at the contribution of the different emitting species to the steady-state intensity, it can be seen that this too is sequence-dependent. With the exception of CC, the largest contribution is from the longest lifetime component which accounts for between 50 and 90 % of the intensity. For CC, the largest contribution to the CC emission spectrum, 51%, comes from the intermediate lifetime component. In all cases, the A-factors are essentially independent of emission wavelength over the range 410 to 450 nm, indicating that the emitting species have very similar emission spectra over this range.

The time-resolved fluorescence decay parameters of pA in double-stranded oligonucleotides are presented in Table 6-8.

		τ_1 / ns	τ_2 / ns	τ_3 / ns	A_1	A_2	A_3	SS_1	SS_2	SS_3	$<\tau>$ /ns
GA	410	0.80	2.06	6.59	0.21	0.75	0.03	0.09	0.81	0.10	1.93
	430				0.21	0.76	0.04	0.08	0.78	0.13	1.96
	450				0.21	0.75	0.04	0.08	0.78	0.13	1.98
GG	410	0.57	2.32	4.86	0.17	0.73	0.09	0.04	0.74	0.21	2.26
	430				0.17	0.73	0.09	0.04	0.76	0.20	2.25
	450				0.16	0.74	0.09	0.04	0.76	0.02	2.25
AT	410	1.00	2.85	5.62	0.22	0.71	0.07	0.08	0.78	0.13	2.62
	430				0.22	0.72	0.06	0.08	0.79	0.13	2.61
	450				0.23	0.70	0.07	0.09	0.78	0.13	2.60
CC	410	0.55	1.16	-	0.66	0.34	-	0.48	0.52	-	0.76
	430				0.68	0.32	-	0.49	0.51	-	0.75
	450				0.67	0.33	-	0.49	0.51	-	0.75
TT	410	1.16	2.16	-	0.38	0.62	-	0.25	0.75	-	1.78
	430				0.40	0.60	-	0.26	0.74	-	1.76
	450				0.40	0.60	-	0.26	0.74	-	1.76

Table 6-8. Fluorescence decay parameters, fractional contributions to steady-state intensity (SS) and average lifetime ($<\tau>$) for pA in double-stranded oligonucleotides with different combinations of nearest neighbours excited at 390 nm. A-factors (A_i) are shown for an emission wavelength of 430 nm. Global χ^2= 1.08 for GA, χ^2= 1.04 for GG, χ^2= 1.01 for AT, χ^2= 1.11 for CC and χ^2= 1.07 for TT. Error in τ_1 = 5%, Error in τ_2 = 3 % Error in τ_3 = <1 %. Error in A_1 = 2%, Error in A_2 = 1 % Error in A_3 = <1 %.

In sequence contexts GA, GG or AT, pA displays three decay components. The dominant component is the intermediate lifetime (τ_2) accounting for just over 70% of the total

population. The short lifetime component (τ_1) accounts for approximately 20% of the population in all three sequences, while the longest lifetime component (τ_3) amounts to less than 10% of the relative population. In all three sequences, the steady-state spectrum is dominated by the intermediate (2-3 ns) lifetime component. In CC or TT, pA has just two lifetimes, the longest lifetime component is absent. Both components have substantial populations and both contribute to the steady-state intensity.

The lifetimes and A-factors for both single and double-stranded oligonucleotide for an emission wavelength of 430 nm are displayed graphically in Figure 6-12.

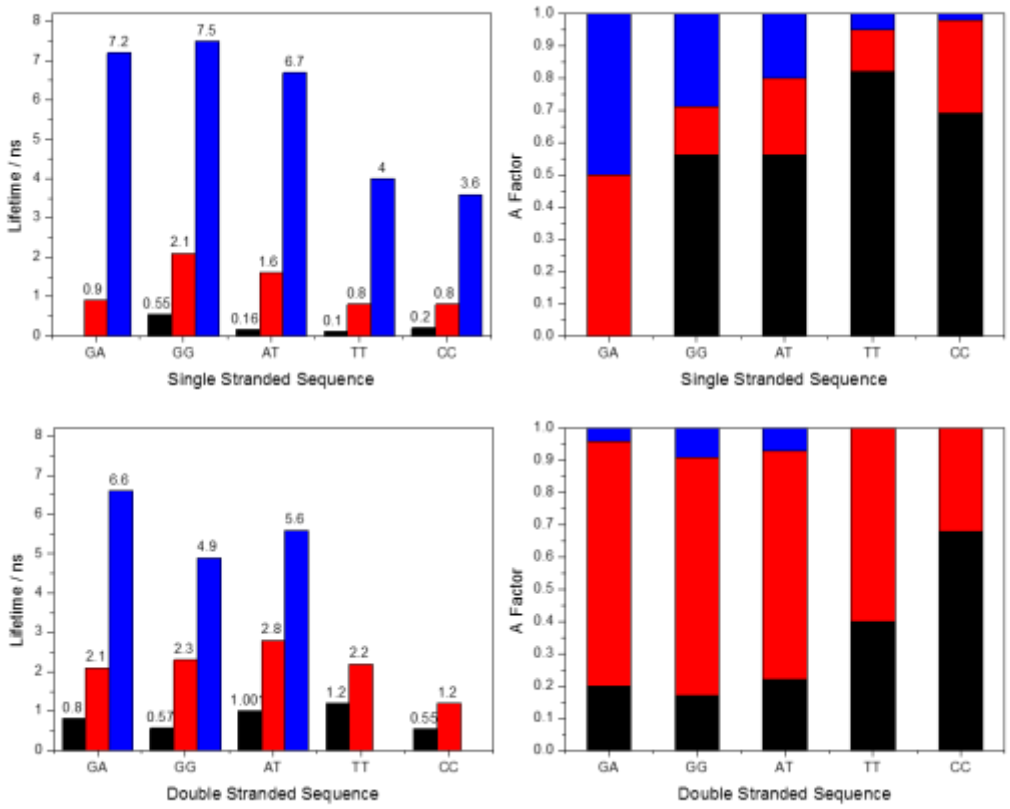


Figure 6-12. Fluorescence decay parameters, for pA in single and double-stranded oligonucleotides with different combinations of nearest neighbours excited at 390 nm. A-factors (A_i) are shown for 430 nm.

6.3.2.3 Comparison of Single- and Double-Stranded Oligonucleotides

There are several differences between the time-resolved decay parameters of pA-containing single and double-stranded oligonucleotides and these are highly dependent on the nature of the neighbouring bases. The number of lifetime components increases from two to three when pA is stacked between G and A, while for sequences GG and AT three lifetime components are sufficient for both single and double-stranded oligonucleotides. On the other hand,

sequences CC and TT require three lifetime components for pA in single-stranded oligonucleotides but only two in double-stranded oligonucleotides.

Although the number of lifetimes is sequence-dependent, some general trends can also be elucidated. In a duplex, the longest lifetime is shorter than the corresponding lifetime in a single strand and also accounts for a smaller proportion of the total population. Where it exists, the intermediate component is longer in the duplex than in the single strands. For all sequences but CC, this intermediate component accounts for the major component of the duplex population.

While these global trends can be resolved in the lifetimes, the overall effect of duplex formation on the average lifetime is different for each sequence. The average lifetime decreases for pA in the oligonucleotide sequences GA and GG on going from a single to a double strand, while in sequences AT, CC and TT, pA displays duplex-enhanced florescence with the average fluorescent lifetime increasing in a double strand. This enhancement has also been observed in the quantum yields (Table 6-3).

It is evident that in single-strand CC and TT the fluorescence of pA is highly quenched by interaction with the neighbouring bases. This is likely due to inter-base charge transfer between the natural base and excited pA. This known to be the major mechanism of inter-base quenching of 2-aminopurine fluorescence in DNA.[10] The quenching efficiency depends on the difference in redox potential between the natural base and the excited state of the base analogue, and the local stacking geometry. For both CC and TT there is a marked decrease in quenching in the double strand; this is particularly great for TT. This suggests that the constraint on base mobility imposed by base pairing inhibits pA from accessing the conformations that are most conducive to charge transfer. Conversely, in the case of GA, which displays a large decrease in average lifetime in the double strand, duplex formation confers a stacking geometry with the neighbouring bases in which pA is more efficiently quenched.

This influence of base pairing on stacking interactions differs significantly from the behaviour of 2-aminopurine, which appears to adopt similar stacking geometries in single strands and double strands (similar lifetime components), with relatively small changes in their populations (A factors) between the two.[10] Unlike 2AP, the structure of pA differs significantly from the natural bases and this will have an effect on stacking interactions.

6.3.2.4 Comparison between pA Riboside and pA in Oligonucleotides

The pA riboside exhibits a bi-exponential decay when free in solution (Table 6-5). As the only difference between the two suggested tautomers is the position of a hydrogen on the Watson-crick hydrogen bonding face (Figure 6-11), it is possible that both tautomers could persist in a single-stranded oligonucleotide. On the basis of the measured decay parameters it is not possible to determine whether this is the case, since the conformational heterogeneity of the oligonucleotides results in multi-exponential decays, even when the free FBA has a single lifetime (as in the case of 2AP).

Where the same species is present in both solution and an oligonucleotide, a linear relationship between average lifetime and quantum yield should exist for the situation where the radiative rate remains constant and there is only a change in non-radiative decay rate due to quenching. Figure 6-13 illustrates this relationship for pA in PBS and pA-containing single-stranded oligonucleotides.

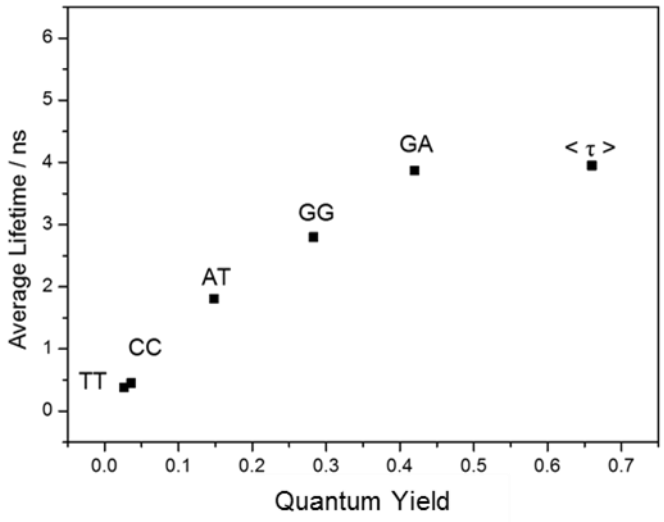


Figure 6-13. Average lifetime vs quantum yield for pA-containing single-stranded oligonucleotides with nearest pA neighbours, GA, GG, AT, CC and TT. < *τ* > is the average lifetime of pA in PBS.

It can be seen that there is a significant discrepancy between the average lifetime of free pA and the trend shown by the average lifetimes of the oligos. There are several factors that could be responsible for, or contribute to, this discrepancy. Firstly, the radiative lifetime of pA could be different in the oligos. However, as discussed below, this is only a small effect. Secondly, it is important to note that the quantum yields were measured at an excitation wavelength of 355 nm, whereas the lifetimes were measured at 390 nm. If there were a significant dependence of the quantum yield (average lifetime) of pA on excitation wavelength, this might account for the discrepancy. This is plausible, given that both decay components contribute significantly

to the average lifetime and they may have different excitation spectra. Finally in the oligos pA may have lifetime components that are too short to be resolved by the present experiments, so-called dark states. (This phenomenon is observed for 2AP in DNA). For example, this could result from the already short, 0.19 ns, lifetime component present in free pA becoming quenched to the extent that its lifetime is less than 30 ps, beyond the detection capabilities of these experiments, in the oligos.

The longest lifetimes observed for pA in certain oligonucleotide sequences are longer than either of the lifetimes exhibited by pA riboside in buffer or ethanol. For pA with nearest neighbours GA, GG and AT these lifetimes are 7.24 ns, 7.48 ns and 6.72 ns respectively. The longest component observed for the pA riboside in PBS is 5.97 ns. This could be due to an increase in the radiative lifetime of pA as a result of incorporation in the oligo. There is a decrease in the molar absorption coefficient by 10-20%, depending on sequence context, when pA is incorporated in the oligos, and there will be a corresponding increase in the radiative lifetime.[21] It is possible that the change in the polarity of the pA environment will affect the non-radiative decay rate. Although the decay parameters have not been measured in any non-polar solvents, the quantum yield is known to be higher in less polar solvents, 0.74 in DCM. This suggests that the fluorescence lifetime in a less polar environment would also be longer. It is also possible that, much like the qAN compounds discussed in Chapter 4, excited state tautomerism is involved in excited state decay and that the more restricted environment of an oligonucleotide could hinder this process. This discrepancy between the lifetimes, together with the resemblance of the oligos emission spectra to that of pA in ethanol, rather than water, indicates that the longest lifetime component is not due to a destacked conformation in which pA is exposed to the aqueous environment, but corresponds to a conformation in which pA is stacked within the strand but is not subject to inter-base quenching.

6.3.2.5 Lifetime Assignment

The multiple lifetimes exhibited by 2AP and 6-MI in oligonucleotides are assigned to differently stacked environments of the FBA. The shortest lifetime is associated with the most well-stacked arrangement, intermediate lifetimes to intermediately stacked environments and the longest lifetime to the least stacked arrangement. For pA, it is likely the multiple lifetimes pA exhibited in the oligonucleotides are also due to multiple stacked conformations, but the interpretation is complicated by the bi-exponential decay of the free base and the fact that there is no obvious relation between these lifetimes and those seen in the oligos. Unlike 2AP, the

lifetimes and the A-factors of pA show large variations depending on the sequence context and whether the oligonucleotide is in a single or a double-stranded form.

In the duplex structures the shortest lifetime component does not account for the largest population. Melting temperature results recorded by Wilhelmsson *et al.* (Table 6-9) indicate that all oligos are strongly hydrogen bonded. The largest pA population would therefore be expected to be pA in a B-form, base-paired type arrangement, implying that a conformation other than the well-stacked, Watson-Crick based paired conformation is the most quenched. Consequently, this result can help to explain the duplex enhanced fluorescence observed in certain sequences. In a single strand, without the additional structure of the duplex, pA is free to take up any stacked conformation. The conformation it adopts preferentially is a more quenched conformation than the Watson-Crick conformation. In a duplex, the additional constraint of base pairing means pA adopts a typical B-form but less quenched position, resulting in a higher quantum yield.

Sequence name	DNA sequence[a]	T_m^{pA} /°C	T_m^{A} /°C)	ΔT_m /°C
AT	5'-d(CGCA**A(pA)T**TCG)-3'	44.8 ±0.1	43.4 ±0.1	1.4 ±0.1
CC	5'-d(CGCA**C(pA)C**TCG)-3'	54.6 ±0.1	50.3 ±0.1	4.3 ±0.2
GA	5'-d(CGCA**G(pA)A**TCG)-3'	43.0 ±0.2	45.3 ±0.2	-2.3 ±0.3
GG	5'-d(CGCA**G(pA)G**TCG)-3'	47.6 ±0.1	48.1 ±0.2	-0.5 ±0.2
TT	5'-d(CGCA**T(pA)T**TCG)-3'	45.8 ±0.3	40.6 ±0.1	5.2 ±0.3

Table 6-9. Wilhelmsson *et al.* melting temperatures of pA-modified duplexes (T_m^{pA}), unmodified duplexes (T_m^{A}) and the difference (ΔT_m) between them. [a] Unmodified sequences contain an adenine instead of pA. Duplexes were formed by annealing with the complementary strand as described in the experimental section.

6.4 Conclusions and Future Work

Pentacyclic adenine has been characterised, free in solution and in a series of single-strand and double-strand oligonucleotides. Both steady-state and time-resolved measurements illustrate the sensitivity of the fluorescence properties of pA to interactions with the natural bases. Time-resolved fluorescence measurements have indicated that pA has two lifetimes in solution and it is likely these originate from two tautomers. In oligonucleotides pA has multiple lifetimes, either two or three depending on the identity of the nearest neighbours. It is possible that similar to the interpretation of 2AP and 6MI lifetimes in oligonucleotides the different lifetimes of pA can be assigned to different stacked conformations. Unlike 2AP and 6-MI, in a duplex the major conformation does not correspond to the most highly quenched, but to an intermediate lifetime with some of the population in a more quenched environment and a few percent of the population in a less quenched environment.

Compared to 2AP, the lifetime behaviour of pA is complex and difficult to interpret. One reason for this could be the relatively large size of pA. 2AP is a structural analogue of adenine whereas pA is larger than adenine by the addition of three 6-membered rings. 2AP can adopt multiple stacked conformers without affecting the global structure of the oligonucleotide, mimicking the natural base, adenine, it can also undergo base flipping and be recognised as a substrate by enzymes. The larger size of pA may make it less able to act in the same manner as a natural base. While its planar, π-rich structure enables it to stack with neighbouring bases in the duplex, it is probable that it does not take up the same stacked conformations as the natural bases and this makes its behaviour less intuitive to understand. Nonetheless, it has been shown that it is possible, to a certain extent, to rationalise the nature of the pA lifetimes and FBAs with complex behaviour, such as 6-MI, have still found use in the study of DNA structure and dynamics.

The high brightness of pA in certain sequence contexts could also make pA a useful probe. The large difference in quantum yield (average lifetime) of pA between single and double strands could make it particularly useful as a probe of base pairing/unpairing and duplex unwinding.

To further understand the inter-base stacking properties of pA it would be useful to run MD simulations, similar to those performed for 6-MI. This should help to establish the nature of conformations that give rise to the various pA lifetime components in oligonucleotides. For 6-MI it was found that the quantum yield was affected not just by the direct neighbour base but also the base in n+2 position. Therefore, widening the investigation of pA-containing sequences could result in identification of a sequence with an even higher quantum yield.

It has been shown that energy transfer occurs from the natural bases to pA. This could make pA a useful probe of energy transfer in DNA. As the quantum yield of pA is higher than that of 2AP, a lower concentration of oligonucleotides would be needed in such experiments. This would allow longer oligonucleotides to be investigated, reducing the problem of the inner filter effect (caused by the intense absorption of the natural bases) that is encountered when using 2AP. Since it has been shown, using time-resolved fluorescence measurements, that energy transfer from the natural bases to 2AP is conformationally selective, favouring highly stacked conformations.[20] Similar energy transfer measurements on pA-containing oligos would shed light on which conformations are most involved in energy transfer, and hence assist in the assignment of fluorescence lifetimes to conformational states.

6.5 References

1 A. M. Michelson, *Nature*, 1958, **182**, 1502.

2 D. Voet, W. B. Gratzer, R. a Cox and P. Doty, *Biopolymers*, 1963, **1**, 193–208.

3 D. Markovitsi, *Photochem. Photobiol.*, 2016, **92**, 45–51.

4 T. Takaya, C. Su, K. de La Harpe, C. E. Crespo-Hernandez and B. Kohler, *Proc. Natl. Acad. Sci.*, 2008, **105**, 10285–10290.

5 M. Kasha, H. R. Rawls and M. Ashraf El-Bayoumi, *Pure Appl. Chem.*, 1965, **11**, 371–392.

6 J. M. L. Pecourt, J. Peon and B. Kohler, *J. Am. Chem. Soc.*, 2000, **122**, 9348–9349.

7 C. T. Middleton, K. de La Harpe, C. Su, Y. K. Law, C. E. Crespo-Hernández and B. Kohler, *Annu. Rev. Phys. Chem.*, 2009, **60**, 217–239.

8 E. R. Bittner, *J. Chem. Phys.*,2006, **125**, 094909.

9 D. G. Xu and T. M. Nordlund, *Biophys. J.*, 2000, **78**, 1042–1058.

10 A. C. Jones and R. K. Neely, *Q. Rev. Biophys.*, 2015, **2**, 1–36.

11 O. F. a. Larsen, I. H. M. van Stokkum, F. L. de Weerd, M. Vengris, C. T. Aravindakumar, R. van Grondelle, N. E. Geacintov and H. van Amerongen, *Phys. Chem. Chem. Phys.*, 2004, **6**, 154–160.

12 M. Narayanan, G. Kodali, V. Singh, Y. Xing, M. E. Hawkins and R. J. Stanley, *J. Phys. Chem. B*, 2011, **114**, 5953.

13 T. M. Nordlund, S. Andersson, L. Nilsson, R. Rigler, A. Gräslund and L. W. McLaughlin, *Biochemistry*, 1989, **28**, 9095–9103.

14 C. R. Guest, R. A. Hochstrasser, L. C. Sowers and D. P. Millar, *Biochemistry*, 1991, **30**, 3271–3279.

15 M. E. Hawkins, W. Pfleiderer, F. M. Balis, D. Porter and J. R. Knutson, *Anal. Biochem.*, 1997, **244**, 86–95.

16 M. E. Hawkins, *Fluorescent pteridine probes for nucleic acid analysis.*, Elsevier Inc., 1st edn., 2008, vol. 450.

17 A. Moreno, J. L. Knee and I. Mukerji, *Biochemistry*, 2012, **51**, 6847–6859.

18 A. Moreno, J. L. Knee and I. Mukerji, *J. Phys. Chem. B*, 2016, **120**, 12232–12248.

19 T. M. Nordlund, D. Xu and K. O. Evans, *Biochemistry*, 1993, **32**, 12090–12095.

20 G. McKenzie, PhD Thesis, The University of Edinburgh, University of Melbourne, 2017.

21 M. Bood, A. F. Füchtbauer, M. S. Wranne, J. J. Ro, S. Sarangamath, A. H. El-Sagheer, D. Rupert, R. S. Fisher, S. Magennis, F. Höök, A. C. Jones, T. Brown, B. H. Kim, A. Dahlen, M. Wilhelmsson and M. Grotli, *Chem. Sci.*, 2018, **9**, 3494–3502

22 P. Sandin, K. Börjesson, H. Li, J. Mårtensson, T. Brown, L. M. Wilhelmsson and B. Albinsson, *Nucleic Acids Res.*, 2008, **36**, 157–67.

23 S. JB and O. HE, *Ind. Eng. Chem.*, 1951, **43**, 2117.

24 J. F. Swindells, C. F. Snyder, R. C. Hardy and P. E. Golden, *Suppl. to Natl. Bur. Stand. Circ.*,1958, **440**, 1–7.

Chapter 7 : Two-Photon Excitation of Fluorescent Base Analogues

7.1 Introduction

Two-photon excitation (TPE) was first postulated by Maria Göppert-Mayer in her doctoral thesis in 1931.[1] As a high photon flux is required for two-photon excitation it was not until the advent of the laser in the 1960s that it could be proved experimentally.[2] TPE did not find widespread use until the 1990s when Denk *et al*. pioneered two-photon laser scanning microscopy.[3] Today two-photon microscopy is a commonly used technique, extensively employed in biological imaging.

Unlike one-photon absorption, where the probability of excitation is directly proportional to intensity, two-photon excitation is proportional to intensity squared. A high enough photon flux to generate excitation will therefore only occur at the focal point of the excitation laser beam within the sample, resulting in highly localised excitation. This small excitation volume, typically a few femtolitres, has a number of advantages.[4]

One advantage of the small focal volume is the high degree of spatial selectivity. A small volume of excitation means fluorescence is only emanating from one point, this reduces out of focus blur and leads to better image contrast.[5] The small focal volume, has a further advantage: a reduction in out-of-focus photobleaching.[6] A comparison can be made between two-photon microscopy and confocal microscopy. Confocal microscopy is a one-photon excitation technique which detects fluorescence emission from a volume similar to that of two-photon microscopy. In a confocal measurement, the laser used as an excitation source will cause excitation along the entire path of the beam, although, in order to have high spatial resolution, detection is restricted to one focal plane along this path. This means the proportion of photobleaching incidents compared to the number of detected photons is high. In a two-photon measurement, excitation occurs only at the focal point, consequently, the proportion of bleaching is much lower. For the same reason photodamage is decreased as the small volume of excited molecules minimises the area of interaction of excited species with the environment.

An additional advantage of TPE is the wavelength of light used, typically near-IR light in the wavelength range 700-1000 nm.[2] Light in the near-IR region is absorbed considerably less by background biological species than light from the UV-visible region.[7] Biological samples will generally contain fluorescent components such as proteins, flavins or porphyrins which absorb in the UV.[8] The reduction in background absorbance therefore reduces autofluorescence

(fluorescent emission emanating from the background). At TPE wavelengths, scattering in tissue is also minimised. There are two predominant types of scattering in biological samples, Mie and Rayleigh scattering.[9] Mie scattering generally refers to scattering from particles of a similar size to the wavelength of light and is largely independent of wavelength. Rayleigh scattering, on the other hand, is scattering from particles significantly smaller than the wavelength of light and is strongly wavelength-dependent. Rayleigh scattering is inversely proportional to wavelength to the fourth power.[8] This makes scattering an order of magnitude less intense in the near-IR than the UV. As well as less scattering in total, the light that does get scattered is less problematic. Scattered two-photon excitation light does not lead to fluorescence. The probability of two scattered photons reaching a target molecule at the same time, in such a manner that they cause excitation is negligible.[10] In a one-photon measurement scattered light can cause excitation, leading to out of focus blur. The combination of reduced scattering and reduced background absorbance allows for deeper tissue penetration to be achieved. The reduction in scattering, minimisation of autofluorescence, and decrease in out-of-focus photobleaching and photodamage also results in a higher signal-to-background ratio.

As well as the many advantages, there are also some disadvantages associated with TPE and these must be weighed against the positives. The main disadvantage is the low two-photon absorption cross-section. As the probability of excitation is lower, light sources with a high power density must be used. The most commonly used excitation source for two-photon microscopy is a mode locked Ti:sapphire laser. This is ideal due to the very short pulses it produces (~ 100 fs). The shorter the pulse, the higher the peak power for a given average power. This maximises photon density while minimising damage caused by heating. A Ti:sapphire laser is also tuneable between 700 to 1000 nm, allowing for a wide range of excitation wavelengths.[6] This type of laser is a more expensive and a more specialised light source than is used in more common forms of microscopy. There can also be damage to cells when using such high intensities. Pigmented samples, for example can suffer photothermal damage due to excessive heating.[11]

Two-photon excitation of fluorescent base analogues (FBA) would be advantageous for all the reasons previously mentioned. There is the additional benefit that fluorescent base analogues tend to have small Stokes shifts. Where excitation and emission wavelengths are similar, in order to stop excitation light being detected, long-pass or band-pass filters are used which often result in part of the emission spectrum being cut off. There is therefore a lowering of the light available for detection. This problem can be alleviated by using two-photon excitation as the excitation wavelength is far removed from the wavelength of detection.

To be of potential use in multiphoton microscopy, fluorescent base analogues must have both a high two-photon absorption cross-section and a high fluorescence quantum yield. The product of the cross-section and quantum yield is referred to as the two-photon brightness and is often a more useful descriptor of the suitability of a fluorophore for two-photon excitation. In order to be widely suitable for use, excitation in the 700 -1000 nm range is also essential.

Several two-photon studies into fluorescent base analogues have been undertaken previously.[12,13,14,15] Quantum yield (ϕ) absorbance (λ_{abs}) and emission maxima (λ_{em}), two-photon cross-section (σ_2), and the two-photon brightness ($\phi\sigma_2$) of several fluorescent base analogues are displayed in Table 7-1. The archetypal base analogue, 2-aminopurine (2AP), has a very low two-photon cross-section, 0.2 GM at 584 nm (1 GM is 10^{-50} cm^4 s photon^{-1}).[12] The brightness value in Table 7-1 (0.13 GM) was calculated using the quantum yield of the free base (0.66). Due to the low fluorescence quantum yield of 2AP in DNA, the brightness in a duplex would be even lower < 0.01.[16] Another disadvantage to two-photon excitation of 2AP is the excitation wavelength, which is around 600 nm and not in the Ti:sapphire wavelength window. Three-photon excitation of 2AP has been achieved and while this involves use of a more desirable wavelength (900 nm) the brightness is still low.

Base Analogue	ϕ	λ_{abs}	λ_{em}	σ_2 / GM	$\phi\ \sigma_2$ / GM
6MAP	0.39	330	430	3.4* (659 nm)[14]	1.3
6MI	0.70	350	430	1.7 (690 nm)[15]	1.2
3-MI (water)	0.88	348	431	1.2 (550 nm)[15]	1.1
5-(thiophen-2-yl)-6-aza-uridine	0.20	332	463	3.8 (690 nm)[13]	0.80
tC	0.21	395	505	1.5 (800 nm)[12]	0.32
2AP (water)	0.66	305	370	0.2 (584 nm)[12]	0.13
5-(thiophen-2-yl)-2'-deoxyuridine	0.01	314	446	7.8 (690 nm)[13]	0.08

Table 7-1. Quantum yield (ϕ), wavelengths of absorption (λ_{abs}) and emission (λ_{em}) maxima, two-photon cross-section (σ_2) and two-photon brightness ($\phi\sigma_2$) of fluorescent base analogues, measured in previous studies. * Cross-section calculated using now obsolete standard values. Correct, recalculated value is lower σ_2 = 2.6 GM.

Another adenine analogue, 6MAP, has the highest two-photon brightness measured so far for a FBA. It has a much higher two-photon cross-section than 2AP, 3.4 GM (at 649 nm).[14] The cross-section of 6MAP was measured by Hawkins *et al.* in 2004, relative to fluorescein which was taken to have a cross-section of 16 GM (690 nm). However, more recent measurements of the fluorescein cross-section report it to be lower (12 GM).[17] Consequently the 6MAP cross-section is lower than the value reported by Hawkins; closer to 2.6 GM. Despite this, even with the lower, recalculated cross-section, the brightness is still one of the highest for any

fluorescent base analogue. 6MAP does, however, share the same disadvantages, of low quantum yield in DNA and unfavourable excitation wavelength, as 2AP.[18]

To date, the highest two-photon cross-section reported for a base analogue is 7.8 GM (at 690 nm), for 5-(thiophen-2-yl)-2'-deoxyuridine.[13] Unfortunately, this is combined with a low quantum yield of only 0.01.[13] In the same study, several nucleosides, each with a slightly different structure, were investigated; these included 5-(thiophen-2-yl)-6-aza-uridine (Table 7-1). It was found that there were large differences in the cross-section, 0.17 to 7.8 GM, despite structural differences being small. This study was significant as it also showed there was no correlation between the two-photon cross-sections and other photophysical properties such as quantum yield.[13]

Of the currently available base analogues, only one, tC, has an absorption maximum above 350 nm. Although this means excitation can be achieved more easily, tC has a low brightness and so is not ideal. The cross-section reported for tC is the only example measured in an oligonucleotide.[12] It was measured in a single-stranded 26-mer, in which tC was flanked by thymine and cytosine (TC). tC is an unresponsive fluorescent base analogue; while most FBAs see their fluorescence quenched on inclusion into an oligo, tC sees almost no effect.[19] A slight increase in the quantum yield was reported when tC was incorporated in this oligo, from 0.20 when free to 0.21.[19] However, the cross-section was not measured for any other sequences, or for free tC, so it cannot be determined if incorporation into an oligo alters the cross-section of tC.

Compared to dyes regularly used for two-photon microscopy, the brightness of even the brightest fluorescent base analogues is low. The two-photon cross-section of rhodamine 6G, at its peak of 690 nm, for example, is 248 GM.[17] Several hundred GM is not uncommon for organic fluorophores and the cross-section of quantum dots can be as high as 10,000's of GM.[20]

There are some general rules for designing fluorophores with high two-photon cross-sections: include long, π-conjugated chains; maximise the distance between electron acceptor and donor, and have a large rigid structure.[21] Due to the constraints already imposed on a successful FBA, small size and base-stacking ability, for example, most of these known principles cannot be directly transferred to FBA design. It is unlikely that brightness as high as several hundred GM will ever be achieved for a fluorescent base analogue, but there is scope to make improvements.

Even without making significant improvement to two-photon cross-sections there is still potential use for FBAs in two-photon microscopy. There are many examples of two-photon imaging in which fluorescent dyes with a cross-section of less than 10 GM have been used. These include fluorophores involved in DNA labelling. For example, 4',6-diamidino-2-phenylindole (DAPI) is a common fluorescent dye used in one-photon and two-photon microscopy. DAPI binds to AT rich regions in DNA, absorbs at 358 nm and emits at 461 nm.[22] It has a two-photon absorbance maximum at around 700 nm with a cross-section of 102 GM.[23] Studies have been carried out using DAPI exciting at longer wavelengths, where it has a lower cross-section. Figure 7-1 shows images of cells labelled with three dyes, DAPI (blue), PATMAN (green) and tetramethylrhodamine (red) and excited at 780 nm; at this wavelength the cross-section of DAPI is just 12 GM.[23,24] There are further examples where DAPI has been excited at wavelengths where the cross-section is even lower. Excitation has been carried out at 840 nm with good results.[25] At this wavelength, the cross-section has not been measured reliably, but based on two-photon spectral shape is likely to be lower than the value of 7.2 GM measured at 820 nm.

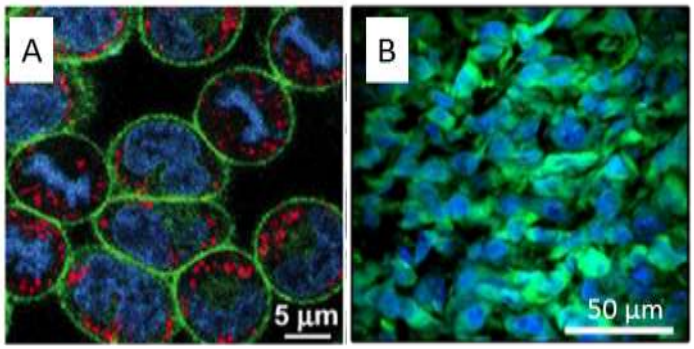


Figure 7-1. Two-photon fluorescence micrographs of cells labelled with three dyes, DAPI (blue), PATMAN (green) and tetramethylrhodamine (red). A) RBL-2H3 cells labelled with 6-diamidino-2-phenylindole (DAPI) excited at 780 nm (Blue).[24] B) GFP-expressing 9L tumor. Cell nuclei stained with DAPI (blue) excitation wavelength 800 nm. [34]

A second fluorescent dye commonly used in fluorescence imaging is ethidium bromide. It intercalates with double-stranded DNA and has emission at around 600 nm.[26] Ethidium bromide has a two-photon cross-section of 7 GM at 870 nm and a quantum yield of 0.13 when bound to DNA and has found use in two-photon microscopy.[27] This is by no means a comprehensive list of fluorescent dyes used in two-photon microscopy but acts to illustrate the use of fluorescent dyes, with low two-photon brightness, when bound to DNA.

The two-photon photophysical properties of commonly used dyes, such as DAPI and ethidium bromide, are not necessarily the same as their one-photon properties. Acker *et al*. compared the one and two-photon absorption and emission spectra of 31 dyes relevant to cell imaging.[28] It was found that there were differences in absorption spectra. Mostly the two-photon

absorption spectrum was blue-shifted, the extent varying from a few nanometres to over a hundred. A red shift was noted in three cases.

It is expected that one and two-photon excitation spectra will be different. In the case of centrosymmetric molecules this is due to difference in selection rules, and for non-centrosymmetric due to vibronic coupling effects. As emission occurs always from the lowest level of the excited state, it is expected that the one and two-photon emission spectra should be identical. In the study by Acker *et al*., however, a red shift of 7-15 nm in the two-photon spectrum was seen was observed for 12 of the dyes. They attributed this to possible calibration errors. In three cases, BODIPY TR, propidium iodide (PI) and aminomethylcoumarin (AMC) the difference was more notable.[28] The one-photon emission spectrum of BODIPY TR has a shoulder, which becomes the dominant peak in the two-photon spectrum. AMC and PI exhibit an emission spectrum with two peaks, on two-photon excitation, but just one peak is observed in their spectra excited by one photon. This could be due to the existence of isomers. If the relative two-photon cross-sections of the isomers differ from their relative one-photon absorption coefficients, the isomer that has greatest emission in the one-photon spectrum may not be the isomer that dominates in the two-photon spectrum.

Although no investigations into the two-photon photophysics of fluorescent base analogues in DNA have been performed, studies into the effect of different protein environments on the two-photon cross-sections other types of fluorophores have been carried out. One of the most well-studied groups is the "Fruit" family of fluorescent proteins. These are a series of mutants, each containing the same chromophore. The electrostatic environment of the chromophore differs between proteins and this results in each protein having a different emission spectrum.[29] The effect of this differing electrostatic environment also has a dramatic effect on the two-photon cross-section of the fluorophore. This is because two-photon cross-section is proportional to the square of the difference between the permanent dipole moments of the initial and final states $|\Delta\mu_{10}|^2$. $\Delta\mu_{10}$ is sensitive to external electric field because of the contribution of the induced dipole moments.[30] The difference in induced dipole moment between the initial and final states originates from the difference in polarisability of the two states. In the fruit proteins, there is a sizeable difference in polarisability due to the charge transfer characteristics associated with the S_0 to S_1 transition. The effect of the local field on the two-photon cross-section is much larger than its effect on the one-photon absorption coefficient. This is because the two-photon absorption cross-section is proportional to both the transition dipole moment squared, $|\mu_{10}|^2$ and the difference between the permanent dipole

moments of the initial and final states, $|\Delta\mu_{10}|^2$, whereas one-photon absorption is proportional only to the former.

An investigation by Goodson *et al.* into dyes binding to DNA found an effect on their two-photon cross-section that can be explained by the same effect.[31] They found that the intercalating dye Hoechst 33258 saw a decrease in two-photon cross-section on binding to DNA whereas groove-binding dye thioflavin saw an increase. They attributed this to the different orientations of the dyes in DNA. It was inferred that Hoechst 33258 would have its dipole aligned more perpendicular to the electric field of DNA whereas thioflavin would be aligned parallel.

Until now fluorescent base analogues have either possessed a low two-photon cross-section, a low quantum yield, an absorbance maximum outside the practical range of two-photon excitation, or a combination of all three. Investigating currently used fluorescent dyes, it can be seen that only slight improvements are required for fluorescent base analogues to be useful in two-photon microscopy. A two-photon cross-section of around 7 GM, as seen with 5-(thiophen-2-yl)-2'-deoxyuridine, combined with a quantum yield of 0.88, as for 3-MI, and an excitation wavelength of 395 nm, as for tC, would be ideal. A number of the recently developed fluorescent base analogues described in Chapters 4, 5 and 6 of this thesis have one-photon absorption spectra in a range that makes them accessible to two-photon excitation using a Ti:sapphire laser, and have relatively high fluorescence quantum yields. The measurement of the two-photon cross-section of several of these fluorescent base analogues will be discussed in this chapter.

7.2 Experimental

7.2.1 Experimental Setup

The experimental setup is illustrated in Figure 7-2 . A mode-locked Ti:sapphire laser (Coherent Mira pumped by Coherent Verdi), producing pulses of duration ~ 150 fs at a repetition rate of 76 MHz, was used as the excitation source. A variable reflective neutral density filter was used to attenuate the excitation beam, which then passed through a dichroic mirror (Semrock Brightline FF735-Di02) and was focused by a 10 × objective (Olympus) into the sample solution, which was contained in a 1 cm path-length cuvette. Fluorescence emission was collected by the same objective, reflected from the dichroic mirror, passed through a short-pass filter (Semrock Brightline FF01-720/SP-25) and detected by a fibre-coupled spectrometer (Ocean Optics USB2000+), with an acquisition time of 2 s. Measurements were made at between twenty and thirty different incident laser powers. Three spectra were collected at each

power and an average taken. The incident power was measured using a Coherent FieldMaster power meter. The sample was dissolved in ethanol (*>99.9%*, Fisher Scientific).

The two-photon cross-sections have an estimated uncertainty of 10% due to uncertainty in the cross-sections of the standards and errors in the measurement of the spectral throughput, absorption spectra and emission spectra.

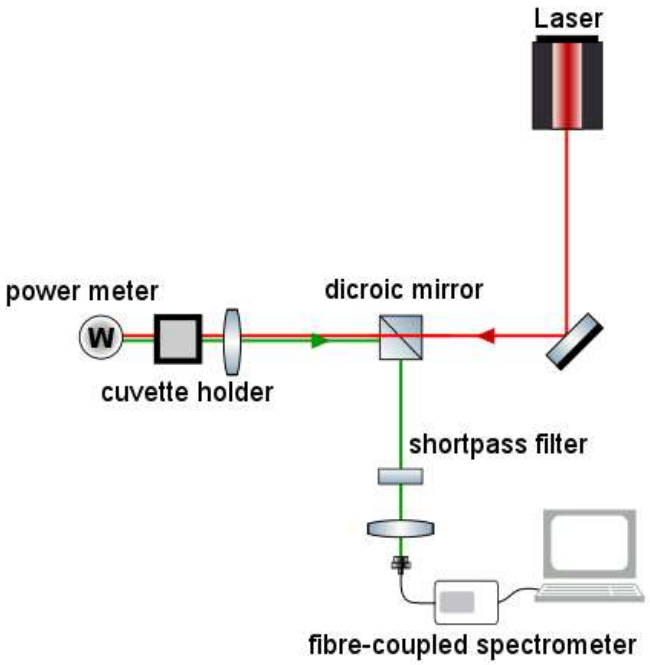


Figure 7-2. Schematic of two-photon experimental system

7.2.2 Two-Photon Cross-section Measurement: Validation of Method

Two-photon cross-sections were measured relative to a standard. Often, reliable values for two-photon cross-sections are not widely available which has led to erroneous cross-section values despite rigorous experimental procedure (for example, in the case of 6MAP, as discussed in section 7.1). To demonstrate that appropriate standards were being used, as well as to validate the experimental system, two-photon absorption cross-sections were measured for three standards, together with the base analogue, pA (free base), the cross-section of which was to be determined.

The three standards chosen were rhodamine 6G in methanol, coumarin 153 in DMSO and coumarin 153 in toluene. These were selected as their cross-sections have been determined, across a wide range of excitation wavelengths, with the purpose of being used as accurate references.[17] The emission spectra of these standards and that of the pA base are shown in Figure 7-3.

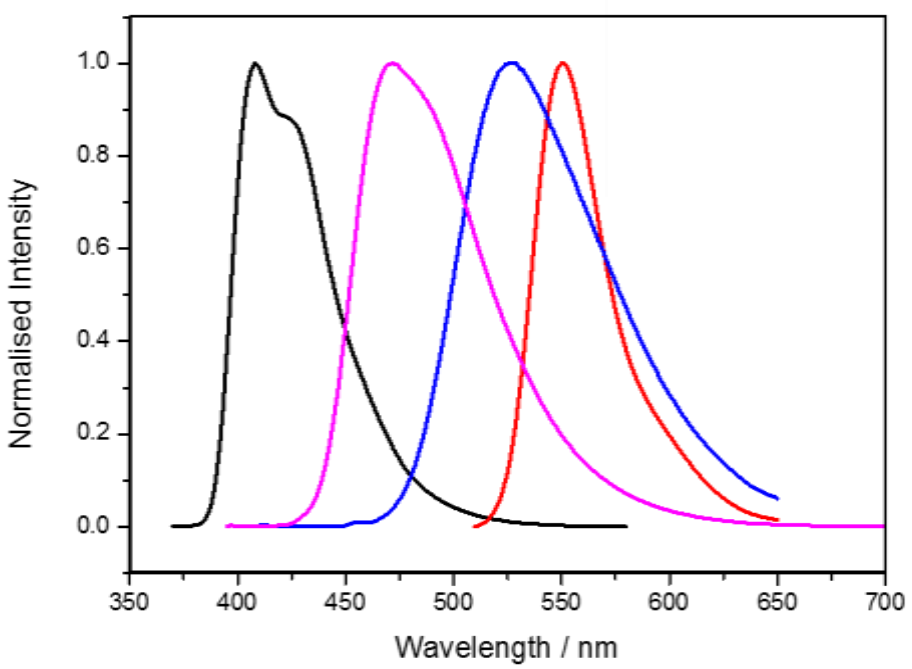


Figure 7-3. Emission spectra of three two-photon cross-section standards, rhodamine 6G in methanol (red), coumarin 153 in DMSO (blue) and coumarin 153 in toluene (pink) and that of the pA base (black).

Two-photon cross-sections (σ_2) were calculated by comparison to a reference using Equation 7-1., the derivation of which was previously described in section 2.2.

$$\frac{\sigma_2^S \phi^S}{\sigma_2^R \phi^R} = \frac{\eta^R n^S c^R F^S \langle P^R \rangle^2}{\eta^S n^R c^S F^R \langle P^S \rangle^2}$$

Equation 7-1.

where ϕ is the quantum yield of fluorescence. η is a term that accounts for the wavelength-dependent collection efficiency of the fluorescence, n is the refractive index of the solvent, c is the concentration, F is the integrated fluorescence signal from the recorded spectrum, P is the excitation power, and the superscripts S and R refer to the sample or reference, respectively.

The two-photon cross-section of the sample can be expressed as Equation 7-2.

$$\sigma_2^S = \frac{\sigma_2^R \phi^R n^S c^R}{\phi^S n^R c^S} \frac{F^S \langle P^R \rangle^2}{\alpha^R F^R \langle P^S \rangle^2}$$

Equation 7-2.

where α^R is correction factor, equivalent to the ratio of $\frac{\eta^R}{\eta^S}$, which corrects for the wavelength-dependence of the detector. The second term can be calculated by measuring the slope of a

plot of integrated emission intensity (F) against power squared ($\langle P \rangle^2$), for sample and reference, respectively, and applying the correction factor.

A description of how each term in Equation 7-2 was determined will be outlined below.

7.2.3 Quantum Yield

The first parameter measured was the quantum yield of each sample, relative to rhodamine 6G in methanol (ϕ = 0.93).[32] For each sample the absorbance and emission intensity of a series of solutions of different concentrations were measured. Figure 7-4 illustrates a plot of integrated fluorescence intensity as a function of absorbance for each sample, where the gradient of each plot is proportional to the quantum yield of that sample. The quantum yields calculated relative to rhodamine 6G are summarised in Table 7-2.

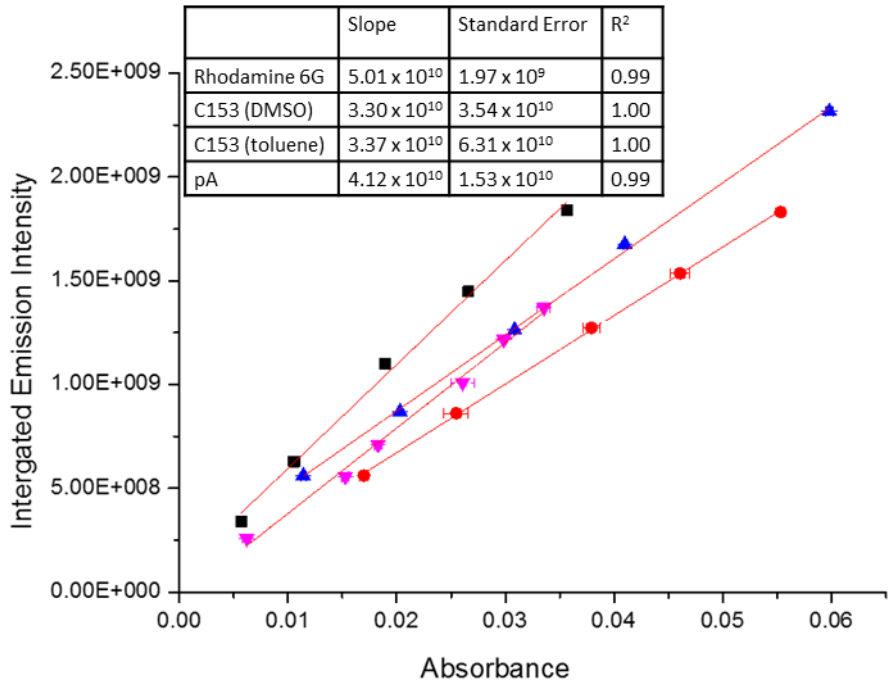


Figure 7-4. Plots of integrated emission intensity vs absorbance for rhodamine 6G in methanol (black), coumarin 153 in DMSO (red), coumarin 153 in toluene (blue) and pA base in ethanol (pink).

	Quantum Yield
Rhodamine 6G in Methanol	0.93 ± 0.04 [32]
Coumarin 153 in DMSO	0.76 ± 0.05
Coumarin 153 in Toluene	0.87 ± 0.04
pA base in ethanol	0.80 ± 0.05

Table 7-2. Quantum yield values determined for coumarin 153 in DMSO, coumarin 153 in toluene and pA base in ethanol measured relative to the standard, rhodamine 6G in methanol.

Literature values for coumarin 153 in DMSO and toluene could not be found. In other solvents Coumarin 153 was found to have a higher quantum yield in less polar solvents, 0.90 in

cyclohexane, 0.38 in ethanol.[33] As the polarity of cyclohexane and toluene are similar, ε= 2.0 and ε= 2.3 respectively, similar values of quantum yield could be expected. The value of 0.87± 0.04 in toluene is similar to the literature value in cyclohexane indicating that the quantum yield measured here is of an expected magnitude.

7.2.4 Spectral Correction Factor

The second parameter to be measured was a spectral correction factor. The contribution to this from the objective, emission filter and dichroic was negligible, as their transmission is flat across the wavelength region of interest. However, the Ocean Optics 2000+ spectrometer has a wavelength-dependent response that derives in part from the spectral characteristics of its detector (Figure 7-5). The wavelength response of the entire spectrometer assembly was not known, therefore the necessary correction factors had to be determined. Correction factors were obtained by measuring the emission spectra on a spectrometer with known wavelength response the Fluoromax 3P spectrometer, and comparing these to the same spectra recorded on the Ocean Optics spectrometer.

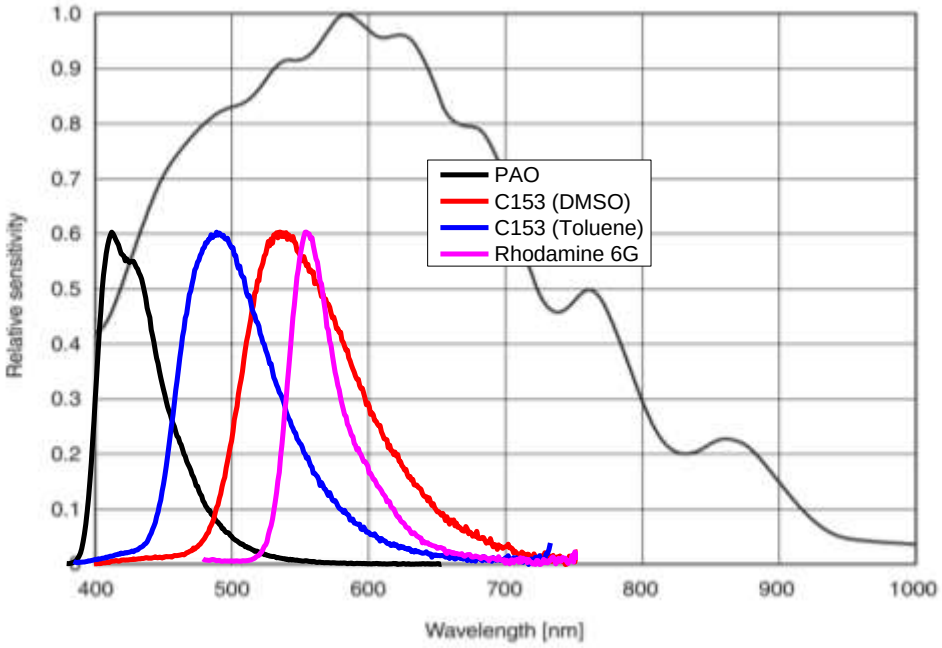


Figure 7-5. Spectral sensitivity characteristics of a SONY ILX511B CCD, (the detector used in the Ocean Optics 2000+ spectrometer) overlayed with the emission spectra of pA in ethanol, coumarin 153 in both DMSO and toluene and rhodamine 6G in methanol.

The same excitation source, the ozone-free xenon arc lamp of the Fluoromax 3P, was used in both cases in order that differences in relative emission could be attributed to detector response. This excitation light was coupled to a cuvette holder using an optical fibre. The emitted fluorescence, from a sample contained within a 1-cm path length fluorescence cuvette,

was collected at 90 degrees to excitation and detected using either the ocean optics or Fluoromax detector.

The Ocean Optics 2000+ contains a SONY ILX511B 2048-pixel CCD with the spectral sensitivity characteristics shown in Figure 7-5. Although this spectral response profile may be somewhat altered by incorporation into the spectrometer it serves as an indication of how significant a correction factor may be expected. As shown in Figure 7-5 the relative sensitivity at 400 nm is only 40% of that at 600 nm. Ideally a standard, emitting over the same range as pA would be used, however, due to the limited number of reliable standards available, this was not possible. In order to obtain accurate results for pA it is imperative to include a correction factor.

Figure 7-6 shows the emission spectra of pA in ethanol, coumarin 153 in both DMSO and toluene and rhodamine 6G in methanol recorded on the Fluoromax-3P and corrected for its wavelength response, and those recorded on the Ocean Optics 2000+ spectrometer. The same sample and excitation source were used for both measurements, only the detector was changed.

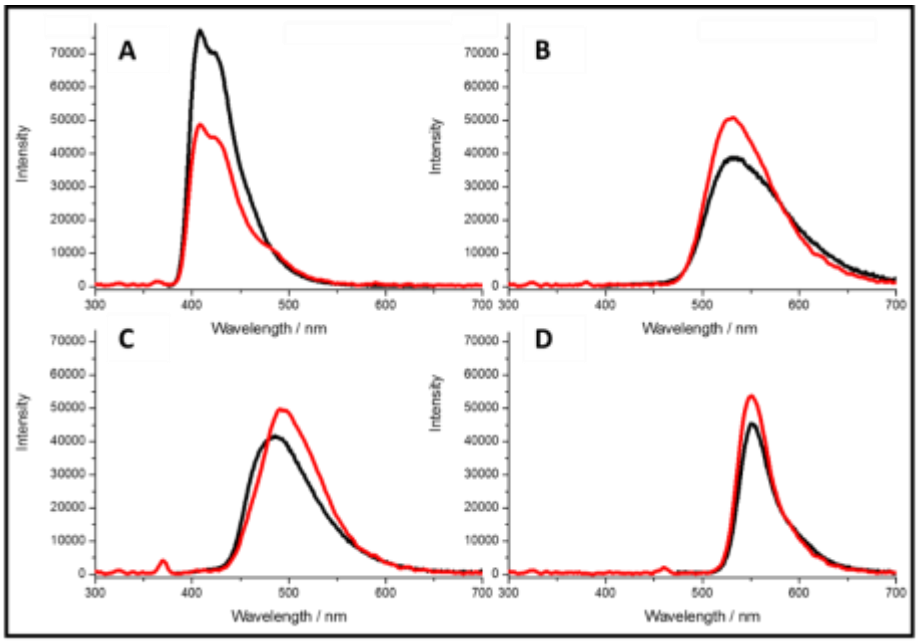


Figure 7-6. Emission spectra of pA in ethanol (A), coumarin 153 in both DMSO (B) and toluene (C) and rhodamine 6G in methanol (D) recorded using a the Fluoromax 3P spectrometer (black) and the Ocean Optics 2000+ spectrometer (red).

Consistent with the spectral response curve shown in Figure 7-3, the relative intensity of the emission spectrum of pA has been underestimated by the Ocean Optics detector which can be seen qualitatively in Figure 7-6. pA has an emission intensity of ~8000 counts per second when measured on the Fluoromax detector but only 5000 when measured on the Ocean Optics detector. In the case of the other three samples, a higher count rate is observed when measured

on the Ocean Optics spectrometer. The total integrated emission values are presented in Table 7-3 and can be used to quantitatively determine the degree of correction required to the intensity ratio of sample and reference standard.

	Integrated Corrected Spectra	**Integrated Ocean Optics Spectra**
Rhodamine 6G (Methanol)	2.34×10^6	2.64×10^6
Coumarin 153 (DMSO)	4.05×10^6	4.48×10^6
Coumarin 153 (Toluene)	3.66×10^6	4.05×10^6
pA (Ethanol)	4.52×10^6	3.14×10^6

Table 7-3. Total integrated area of corrected spectra collected using the Fluoromax 3P spectrometer and uncorrected spectra collected using the Ocean Optics spectrometer.

The two-photon cross-section of a sample was calculated relative to a standard. In some instances rhodamine 6G was treated as a sample with coumarin 153 as the reference in other cases the roles were reversed. For each sample: reference pair, a specific correction factor was calculated, using Equation 7-3:

$$\alpha^R = \frac{C_{reference} O_{sample}}{C_{sample} O_{reference}}$$

Equation 7-3.

where α^R is the correction factor that must be applied to the reference, C is the integrated emission intensity of the corrected emission spectrum collected using the Fluoromax and O is the integrated emission intensity of the spectrum collected using the Ocean Optics spectrometer. Reference and sample refer to whether the solution measured is the standard being used or the sample of interest.

Using Equation 7-3. correction factors were calculated for each reference: sample pair. These are shown in Table 7-4.

Reference : Sample	**Correction Factor**
Rhodamine: Coumarin 153 in DMSO	0.979
Rhodamine: Coumarin 153 in toluene	0.982
Coumarin 153 in DMSO: Coumarin 153 in toluene	0.998
Rhodamine: pA	0.616
Coumarin 153 in DMSO: pA	0.629
Coumarin 153 in toluene: pA	0.628

Table 7-4. Correction factors for each reference: sample pair. The reference:sample intensity ratio measured on the Ocean Optics spectrometer is multiplied by the correction factor.

The correction factors presented in Table 7-4 are largely consistent with the relative spectral profile of the SONY ILX511B 2048-pixel CCD. The rhodamine: coumarin pairs require very little correction whereas when rhodamine or coumarin is used as a standard with pA, the intensity ratio must be corrected by a factor of around 0.62.

To test the robustness of this correction technique, samples were excited at several different wavelengths. Provided a high enough count rate could be achieved, the same correction factor could be reliably measured to within an error of ± 0.01. This was also true of measuring multiple concentrations of the same sample.

7.2.5 Power-dependence of Fluorescence Intensity

The third parameter to be measured was the integrated fluorescence intensity as a function of the square of the excitation laser power. For each compound, emission spectra were collected over a range of average laser powers, between 80 and 320 mW. Between twenty and thirty measurements were made for each sample. At each power, three measurements were made and an average value calculated. Errors were calculated from the standard deviation of these three measurements. The results, before application of a correction factor, are shown in Figure 7-7.

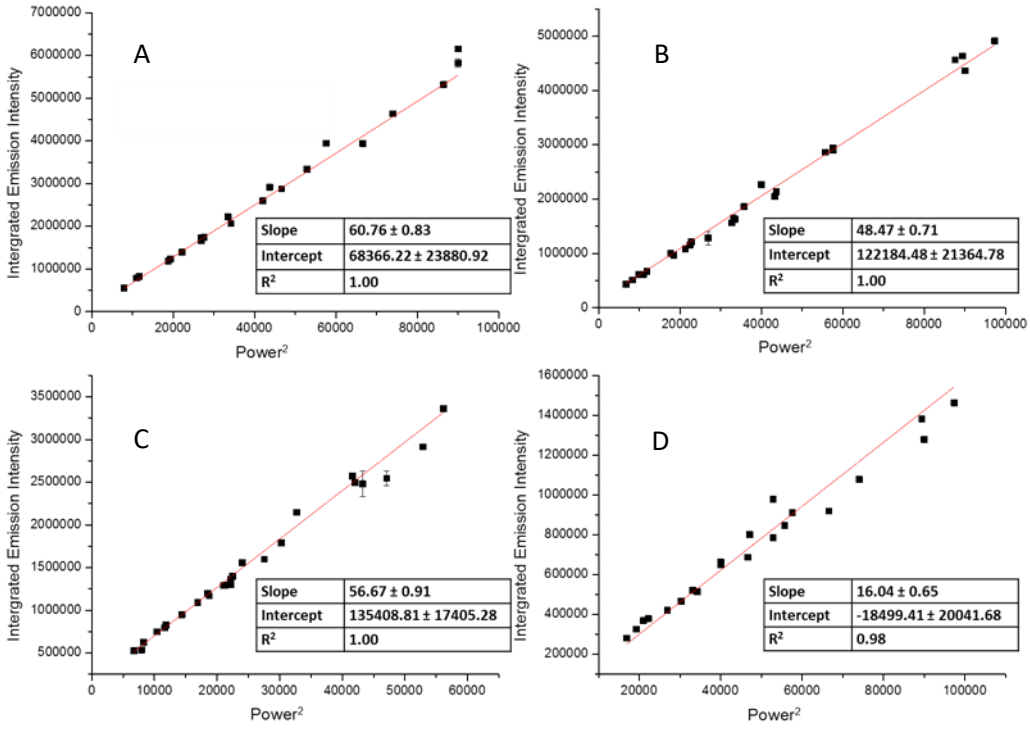


Figure 7-7. Integrated emission intensity as a function of Power² for A) coumarin 153 in DMSO, B) coumarin 153 in toluene, C) rhodamine 6G in methanol and D) pA in ethanol. All collected under the same conditions with excitation at 780 nm.

7.2.6 Calculation of Two-photon Cross-sections of Standards

Based on the uncorrected data shown in Figure 7-7, two-photon cross-sections for the three standards were calculated, as shown in Table 7-5. The spectral correction factor was then applied (Table 7-4), and these results are also shown in Table 7-5. Rhodamine 6G in methanol, coumarin 153 in DMSO and coumarin 153 in toluene have known two-photon cross-sections at 790 nm of 70 GM, 11 GM and 14 GM, respectively.[17]

Sample (Reference)	σ (un-corrected) /GM	σ (corrected) / GM	σ (literature) / GM. [17]
Rhodamine 6G (C153 in DMSO)	68.5	69.9	70
Rhodamine 6G (C153 in toluene)	67.4	68.5	70
C153 in DMSO (Rhodamine 6G)	11.2	11.0	11
C153 in DMSO (C153 in toluene)	10.8	10.8	11
C153 in toluene (Rhodamine 6G)	14.5	14.3	14
C153 in toluene (C153 in DMSO)	14.2	14.2	14

Table 7-5. Two-photon absorption cross-sections of each standard (Coumarin 153 in DMSO, Coumarin 153 in toluene and Rhodamine 6G in methanol), measured relative to each of the other two standards. Values before and after correction for the wavelength response of the spectrometer are shown. Literature values of the cross-sections are given for comparison. Error in the values of σ_2 for the standards is reported as 8%.[17]

For each pair of standards, the two-photon absorption cross-section measured is well within the error margin of the literature values, which is reported at 8%. These standards can, therefore, reliably be used as references to measure the two-photon absorption cross-section of the base analogues.

As it was the first unknown sample investigated, the cross-section of pA was measured relative to all three standards. For subsequent measurements on other fluorescent base analogues, measured at different excitation wavelengths, the cross-section of rhodamine 6G relative to C153 in DMSO was used as validation. Once it had been established that rhodamine 6G and C153 two-photon cross-sections could be reliably cross-correlated, rhodamine 6G alone was used as the standard to measure the FBA cross-sections.

7.3 Results and Discussion

To be potentially useful as probes suitable for two-photon excitation, fluorescent base analogues should have one-photon absorption in the range of 350-500 nm (an indication that the two-photon absorption will be in the 700 to 1000 nm range) and high quantum yields. Of the fluorescent base analogues investigated so far in this thesis, the molecule that fulfils these

criteria most closely is pA, which has an excitation maximum of 391 nm and a quantum yield of at least 0.64 in all solvents.

7.3.1 Two-photon Spectroscopy of pA

The steady-state fluorescence properties of pA were studied following excitation at 780 nm, a wavelength anticipated to allow two-photon excitation. To confirm that absorption was occurring *via* a two-photon process, fluorescence intensity was measured as a function of average laser power. The results are presented in the form of a log-log plot in Figure 7-8. It can be seen that the data are well-fitted by a straight line with a slope of 2.00 ± 0.05 confirming that a two-photon absorption process is being observed

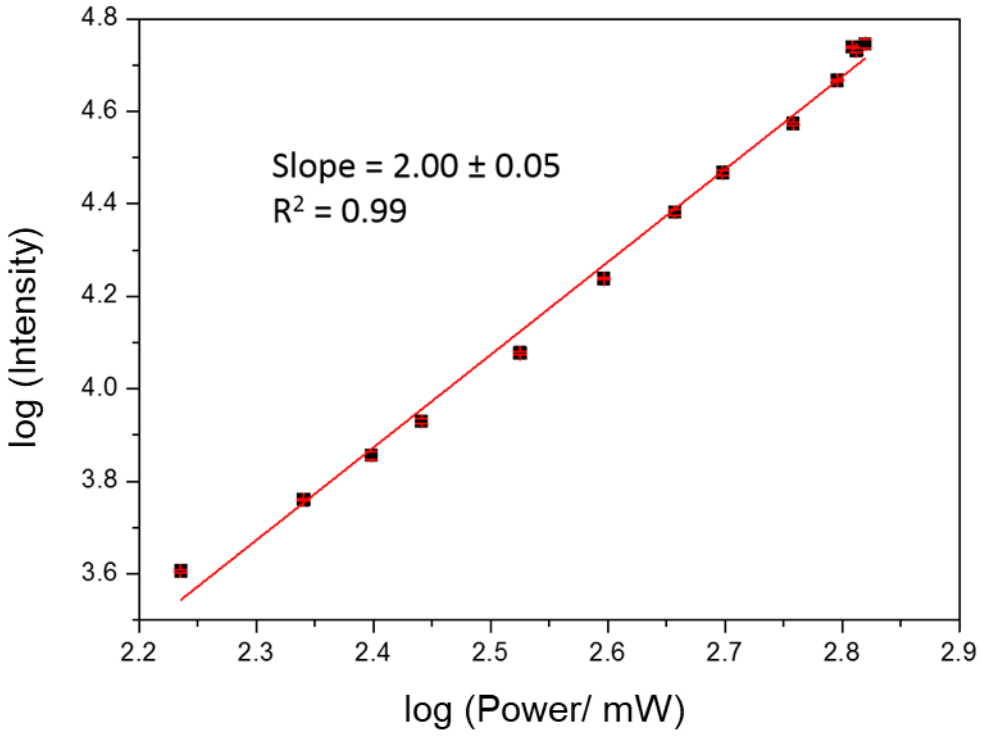


Figure 7-8. Log-log plot of fluorescence intensity vs laser power for pA in ethanol. From the line of best fit, the slope is 2.00 with an R-squared value of 0.99. Excitation was performed at 780 nm, laser power was varied from 660 mW to 172 mW. Error in emission intensity was calculated from the standard deviation of three measurements collected at each power. Error is less than 5%, hence the error bars are smaller than the data-point symbols.

Figure 7-9 shows the emission spectrum resulting from excitation at 780 nm and the corresponding emission spectra resulting from one-photon excitation at 360 nm and 390 nm. Interestingly, two-photon excitation results in a different spectral profile than that seen for one-photon excitation (this is not an artefact of the wavelength response of the ocean optics spectrometer). The shoulder seen in the one-photon emission spectrum dominates in the two-

photon spectrum. This is consistent with the presence of two emitting species, as inferred from the fluorescence lifetime measurements (τ_1= 0.16 ns, A_1= 0.20; τ_2=4.82 ns, A_1= 0.80) which were presented in section 6.3.2.1. These two species, possibly tautomers, appear to have overlapping emission spectra, but differ in their relative 1P and 2P absorption cross-sections. Calculating the steady-state contribution of each of the species it was determined that the one-photon emission spectrum is almost exclusively (99%) due to the longer-lifetime component (τ_2). To establish if a change in the ratio of these two species, when τ_1 accounts for so little of the emission spectrum, could be responsible for the difference in the one-photon and two-photon spectral profiles, the two spectra were normalised at the 430 nm peak. It was found that the difference in spectral shape accounts for just 1% of the total integrated spectrum. It is therefore possible that preferential two-photon excitation of the short-lifetime component could account for the effect on spectral shape.

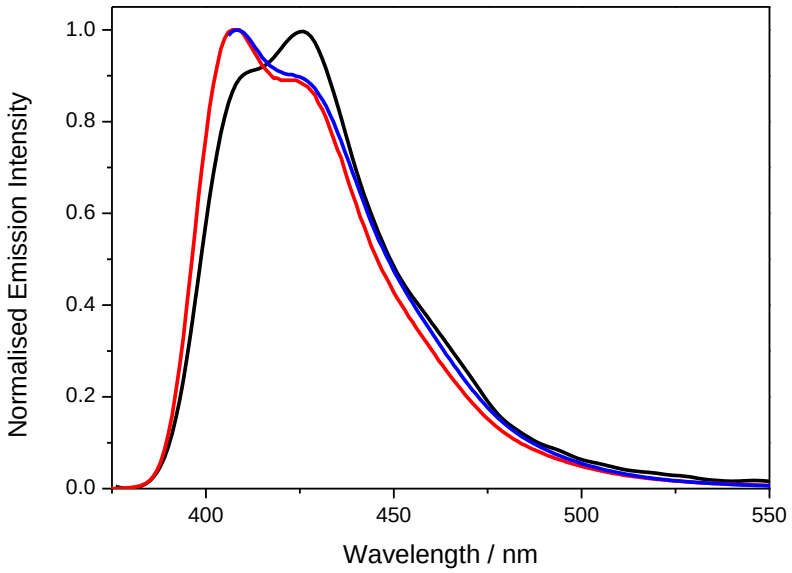


Figure 7-9. Emission spectra of pA in ethanol following one-photon excitation at 360 nm (red) and 390 nm (blue) and two-photon excitation at 780 nm. 2P spectra recorded on ocean optics spectrometer but corrected for wavelength dependent spectral response.

The two-photon excitation spectrum was collected by exciting pA at different wavelengths between 740 nm and 820 nm. The laser power was kept constant at each excitation wavelength by adjusting the neutral density filter (Figure 7-2). The excitation spectrum determined by integrating the emission intensity at each excitation wavelength is shown is shown in Figure 7-10, in comparison with the one-photon excitation spectrum (at an emission wavelength of 420 nm). The two-photon spectrum obtained by measuring the emission intensity at a single emission wavelength of 420 nm is also shown, to insure a valid comparison with the one-photon spectrum.

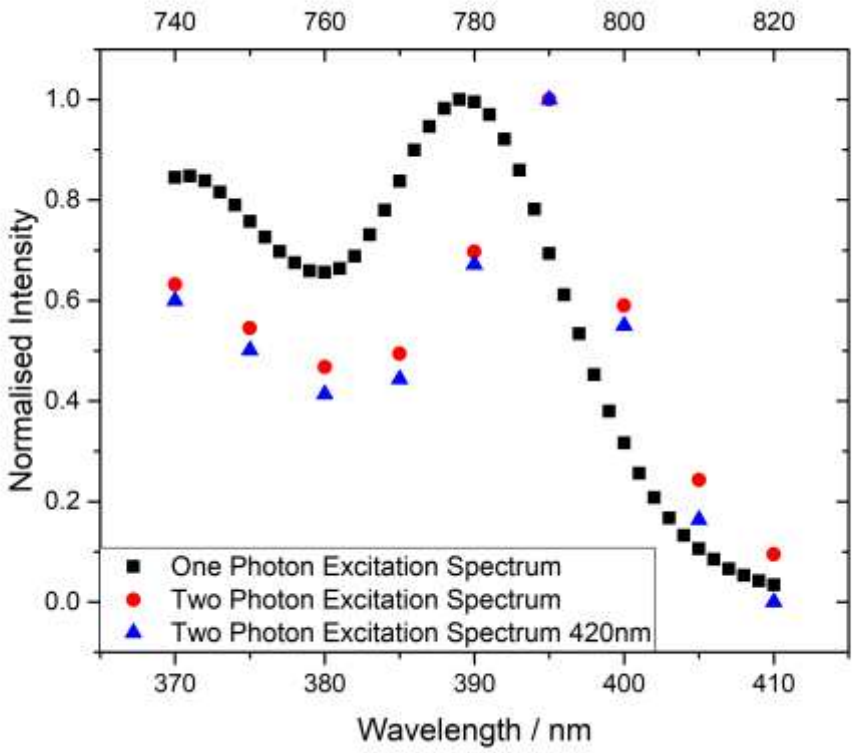


Figure 7-10. Excitation spectra of pA, one-photon excitation (black squares, bottom axis) and two-photon excitation, emission at 420 nm (red circles, top axis).

It can be seen that the two-photon excitation spectrum is red-shifted by around five nanometres compared to the one-photon excitation spectrum, when comparing the peaks at 390/790 nm. Comparing the second maximum, (370/740 nm) the two-photon excitation peak does not appear red-shifted but either occurs at the same wavelength or is blue-shifted. Measurements at shorter two-photon excitation wavelengths are needed to confirm this. The relative intensity of the two peaks is also affected. The 790 nm peak is more intense relative to the 740 peak for two-photon excitation than for one-photon excitation.

It is common to see a blue shifted two-photon excitation spectrum. This is usually attributed to the 0-1 vibronic transition being dominant in two-photon absorption spectrum rather than 0-0 which is most common in one-photon. As a red shift is seen in this case, it is unlikely to be a vibronic effect but is consistent with isomerism, as inferred from the emission spectrum. If the two tautomers contribute differently to the 1P and 2P emission spectra and the excitation spectra of the tautomers are different, a difference in 1P and 2P excitation spectra will be seen. The 1P excitation spectrum varies slightly depending on the wavelength at which emission is recorded, which would indicate that the excitation spectra of the species do differ and can explain the difference observed.

The emission spectra recorded as a function of two-photon excitation wavelength are shown in Figure 7-11.

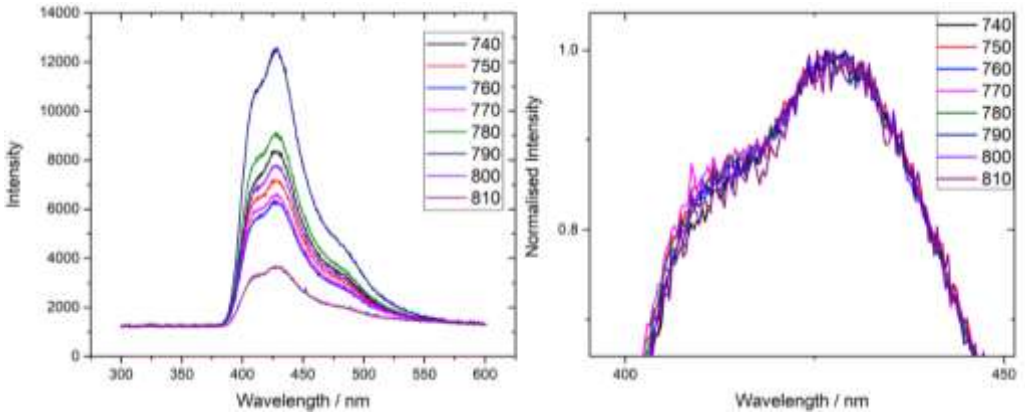


Figure 7-11. pA emission spectra recorded as a function of two-photon excitation wavelength with excitation at 10 nm intervals between 740 nm and 810 nm.

The 2P emission spectra recorded as a function of 2P excitation wavelength do not vary in terms of spectral profile, only intensity. It is possible there is a slight difference in terms of the relative height of the 410 nm and the 430 nm peaks but due to the resolution of the spectra, this cannot be clearly identified. This result is consistent with two species with overlapping emission spectra that differ in their relative 1P and 2P absorption cross-sections.

7.3.2 Two-Photon Cross-section of pA

The two-photon absorption cross-section of pA has been measured relative to three standards: rhodamine 6G in methanol; coumarin 153 in DMSO and coumarin 153 in toluene, as described in section 7.2. The two-photon cross-section of pA calculated relative to each of these standards is reported in Table 7-6.

Reference standard	σ_2 / GM (780 nm)	$\phi\sigma_2$ / GM
Rhodamine 6G in methanol	6.78	5.4
Coumarin 153 in DMSO	6.62	5.3
Coumarin 153 in toluene	6.53	5.2
Average	**6.64 ± 0.5**	**5.3**

Table 7-6. Two-photon cross-section (σ_2) and two-photon brightness ($\phi\sigma_2$) of pA in ethanol, measured relative to three standards, rhodamine 6G in methanol, coumarin 153 in DMSO and coumarin 153 in toluene. Error in the standard values of σ_2 is reported as 8%.

The two-photon cross-section of pA is of a similar magnitude to the highest values previously reported for fluorescent base analogues.[13] Until now FBAs with two-photon absorption cross-sections of this magnitude have had low quantum yields. The high quantum yield of pA combined with a two-photon cross-section of 6.6 GM means the two-photon brightness is the highest reported for any FBA at the time of measurement.

7.3.3 Two-photon Spectroscopy of pA in Oligonucleotides

The two-photon induced emission spectra and the two-photon cross-sections of pA in five oligonucleotides, single and double strands, with varying nearest-neighbour bases were measured. The oligonucleotide sequences studied are shown in Table 7-7. The one-photon photophysical properties of these pA-containing oligonucleotides were discussed in Chapter 6.

Sequence name	DNA sequence
AT	5'-d(CGCA**A(pA)T**TCG)-3'
CC	5'-d(CGCA**C(pA)C**TCG)-3'
GA	5'-d(CGCA**G(pA)A**TCG)-3'
GG	5'-d(CGCA**G(pA)G**TCG)-3'
TT	5'-d(CGCA**T(pA)T**TCG)-3'

Table 7-7. Pentacyclic adenine-containing deoxyoligonucleotides (10mers). Sequences differ only by the neighbouring bases of pA, and it is these neighbouring bases that give their name to the sequence.

pA is environmentally sensitive and its fluorescence is quenched when in oligonucleotides, lowering the quantum yield and thus lowering the two-photon brightness. Therefore, it is important to investigate the effect on the cross-section when pA is incorporated into an oligo.

In each oligo, pA was excited at 780 nm. To demonstrate a two-photon process was taking place a log-log plot of intensity against laser power was plotted. The plots obtained for GA single and double strands are displayed in Figure 7-12. Gradients of ~ 2 in both cases indicate that two-photon absorption is taking place.

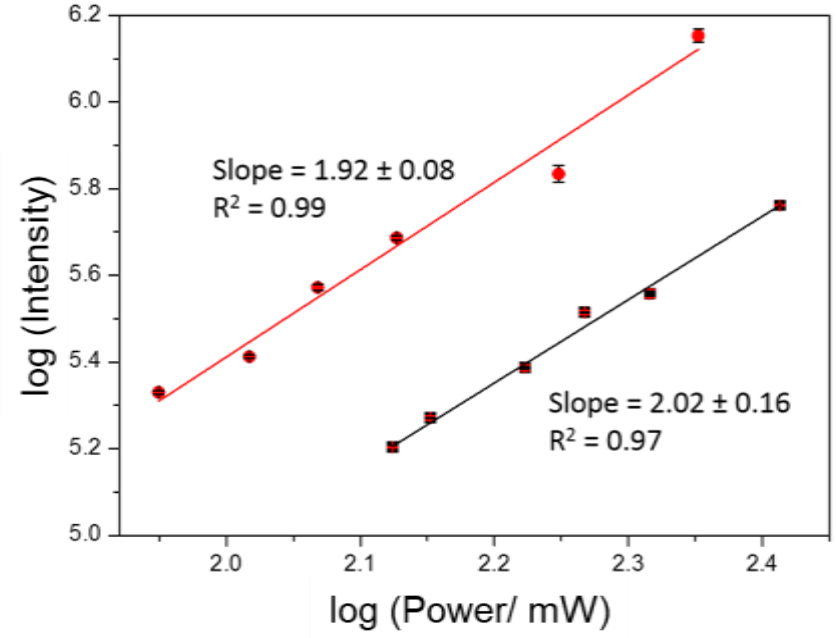


Figure 7-12. Log-log plot of fluorescence intensity vs laser power for the pA–GA oligo, single stranded (red), double-stranded (black). Excitation wavelength was 780 nm. Laser power was varied from 90 mW to 300 mW.

A difference in emission spectral profile for one-photon and two-photon excitation was observed in the oligonucleotides as is shown in Figure 7-13. This is similar to that seen for the free base (Figure 7-9), but the difference is much less pronounced in the oligonucleotides. As with the free base, this could imply the presence of multiple emitting species within the oligonucleotide each having different relative one-photon and two-photon absorption cross-sections. In section 6.3.2.5 the different lifetimes of pA in the oligonucleotide were attributed to different stacking conformations. As the local electrostatic field is known to have an effect on two-photon cross-section in fluorescent proteins (as discussed in section 7.1 above), it is possible that the local base-stacking geometry has an analogous effect on the two-photon cross-section of pA in DNA. Differences in spectral profile could therefore be the result of differently stacked populations having slightly different cross-sections. From the lifetime measurements it was determined that the main contribution to the single-stranded one-photon emission spectrum was the long-lifetime component, the component attributed to pA in the most un-stacked environment. In a double strand, the biggest contribution was from an intermediate lifetime, attributed to a base-paired population. In both single- and double–stranded one-photon emission spectra, the emission peak at 400 nm is dominant over the 430 nm peak; therefore a change in the relative magnitude of the long- and intermediate-lifetime components would not be expected to affect the relative intensity of the two emission spectral peaks. The change in spectral profile upon two-photon excitation should therefore be attributed to a change in the relative two-photon cross-section of the shortest lifetime component.

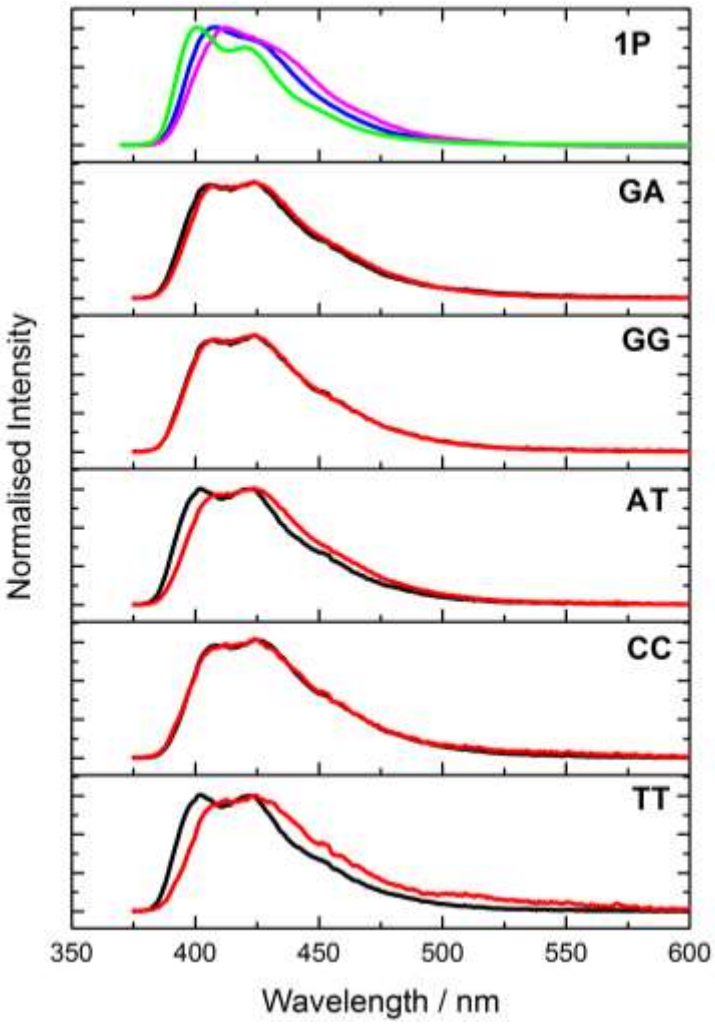


Figure 7-13. Two-photon-excited emission spectra of pA in single stranded DNA (red) and double stranded DNA (black) with different combinations of nearest neighbours in phosphate-buffered saline (PBS), excited at 780 nm. For comparison, the top panel shows the 1P-excited emission spectrum of pA in GA (blue), TT single strand (pink) and TT double strand (green). GA single and double-strands have almost identical spectra. 2P spectra recorded on ocean optics spectrometer but corrected for wavelength response of detector.

7.3.4 Two-Photon Cross-section of pA in Oligonucleotides

The two-photon cross-section of pA in solution is highly promising. Previous two-photon cross-section measurements of environmentally-sensitive fluorescent base analogues have only considered the free base. Investigating the effect of incorporation of the base analogue in DNA is an important next step. The two-photon cross-sections and brightness values of pA in five oligos, both single and double stranded, are presented in Table 7-8.

	Single Stranded		Double Stranded	
	σ_2 / GM	$\phi\sigma_2$ / GM	σ_2 / GM	$\phi\sigma_2$ / GM
GA	2.98	1.25	2.44	0.35
GG	2.29	0.65	2.41	0.30
AT	2.63	0.39	2.68	0.46
CC	2.45	0.09	2.99	0.13
TT	2.49	0.06	2.71	0.30

Table 7-8. Two-photon cross-section (σ_2) and brightness ($\phi\sigma_2$) of pA in single- and double-stranded oligonucleotides with different combinations of nearest neighbours. Cross-sections were measured relative to two standards; rhodamine 6G in methanol and coumarin 153 in DMSO, and average values are presented. Error in the standard values of σ_2 is reported as 8%.

The two-photon cross-section of pA in oligonucleotides is considerably reduced compared with that of the free base. The cross-sections of pA-containing oligos range from 2.98 GM, in GA to 2.29 GM in GG: this is much lower than the value of 6.64 GM for pA free in ethanol. Compared to the difference in molar absorption coefficient, displayed in Table 7-9, the change in two-photon cross-section is considerable. This much larger decrease in the two-photon cross-section than in the one-photon absorption coefficient, is attributed to the two-photon absorption being more sensitive to changes in the local electric field than the one-photon absorption (as discussed in Section 7.1). Cross-section values are similar for single and double strands and do not show any significant dependence on sequence context. The two-photon brightness shows much greater sequence-dependence due to the sensitivity of the quantum yield to the nature of the neighbouring bases.

	Single Stranded					Double Stranded				
	σ_2 / GM (780 nm)	ε_{390} / $M^{-1}cm^{-1}$	ϕ	$\phi\sigma_2$ / GM	$\phi\,\varepsilon_{390}$ / $M^{-1}cm^{-1}$	σ_2 / GM (780 nm)	ε_{390} /$M^{-1}cm^{-1}$	ϕ	$\phi\sigma_2$ / GM	$\phi\,\varepsilon_{390}$ / $M^{-1}cm^{-1}$
GA	2.98	14990	0.42	1.25	6296	2.44	14000	0.14	0.35	1960
GG	2.29	14340	0.28	0.65	4015	2.41	13190	0.13	0.30	1715
AT	2.63	14970	0.15	0.39	2246	2.68	14520	0.17	0.46	2468
CC	2.45	14520	0.036	0.09	523	2.99	12850	0.043	0.13	514
TT	2.49	14110	0.03	0.06	423	2.71	13560	0.11	0.30	1492

Table 7-9. Two-photon cross-sections and brightness compared with one-photon absorption coefficients and brightness, for pA in single- and double-stranded oligonucleotides.

The calculation of two-photon cross-sections reported here is based on the assumption that quantum yield is independent of the route to excitation. Where there are multiple absorbing/emitting populations in a sample, for pA in an oligo these are the result of multiple stacking conformations, the steady-state quantum yield is actually the average quantum yield of the various absorbing species. Each of these pA populations contributes a different extent to the average quantum yield. Should these populations have slightly different relative two-photon cross-sections compared with their relative one-photon absorption coefficients, the

average quantum yield will differ between two-photon and one-photon excitation. There is evidence for this effect in the change in emission spectrum seen between one-photon and two-photon excitation (Figure 7-13), indicating emission from different excited state populations. The difference in the emission spectra is relatively small suggesting this is not having a significant effect. This does, however, introduce a further uncertainty in the calculated cross-section, making it harder to distinguish between genuine differences and uncertainty in the measurement. The cross-section of pA in each oligonucleotide should therefore be considered to be within error of each other.

7.3.5 Two-photon Spectroscopy of qAN1 and qAN4

The one-photon absorption spectra of qAN2 and 3 are not in the ideal range for excitation with a Ti: Sapphire laser (339 nm qAN2, 327 nm qAN3). For qAN1 and qAN4 the excitation maxima are just within the desired range (353 nm qAN1, 356nm qAN4). As the two-photon excitation spectrum of pA was red shifted slightly compared to the one-photon spectrum, it is also possible that this would be seen with qAN 1 and 4, making their two-photon excitation more accessible. Due to the 735 nm dichroic mirror used in the experimental set-up (Figure 7-2), in practice it was not possible to excite at a wavelength below 740 nm. Structures of qAN1 and qAN4 are displayed in Figure 7-14.

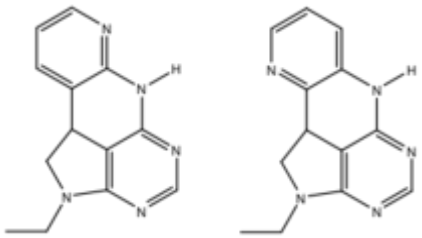


Figure 7-14. Structures of qAN1 and qAN4

Figure 7-15 shows the emission spectra for qAN1 and qAN4 resulting from excitation at 740 nm and the corresponding emission spectra resulting from one-photon excitation at 370 nm. Two-photon excitation results in the same spectral profile as that seen for one-photon excitation. The slight broadening of the qAN1 spectrum is most likely an artefact, caused by applying a correction to the two-photon spectrum (to compensate for the wavelength response of the spectrometer).

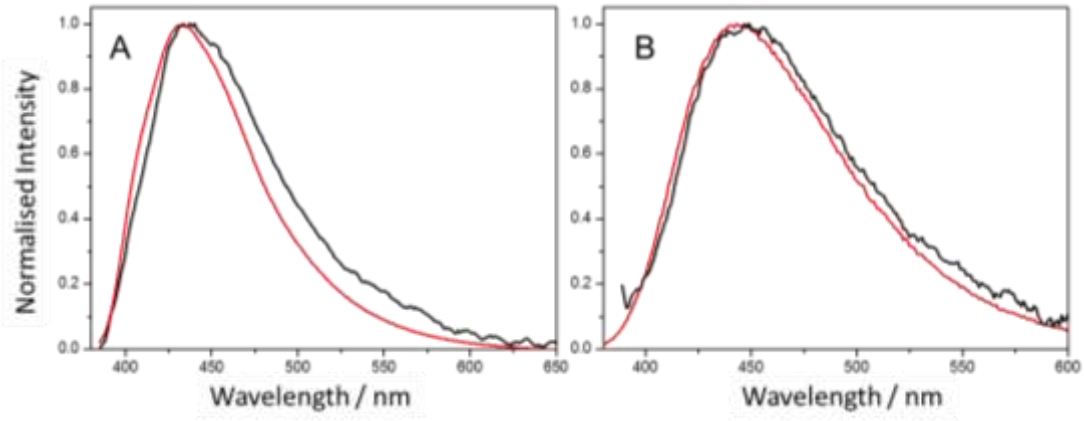


Figure 7-15. Emission spectra of A) qAN1 and B) qAN4 in phosphate buffer following one-photon (red) and two-photon (black) excitation, at 370 nm and 740 nm, respectively.

The two-photon cross-sections of qAN1 and qAN4, measured relative to rhodamine 6G are reported in Table 7-10.

	σ_2 / GM (740 nm)	$\phi\sigma_2$ / GM
qAN1	0.82 ± 0.05	0.15
qAN4	1.03 ± 0.1	0.33

Table 7-10. Two-photon cross-section (σ_2) and two-photon brightness ($\phi\sigma_2$) of qAN1 and qAN4 in phosphate buffer, measured relative to rhodamine 6G in methanol. Error in the standard values of σ_2 is reported as 8%. Error in cross-section was calculated from the standard deviation of seven measurements at various laser powers.

The two-photon cross-sections of qAN1 and qAN4 are considerably lower than the cross-section reported for pA, and the two-photon brightness values are an order of magnitude lower. As the cross-sections have not been recorded at the excitation maximum, it is likely these values are lower than the optimum values. However, as the brightness is an order of magnitude lower than pA, it is unlikely that qAN1 or qAN4 will surpass the brightness of pA, even when excited at the maximum.

7.3.6 Two-photon Spectroscopy of Extended-6-azauridines

The extended 6-azauridines, shown in Figure 7-16 and discussed in Chapter 4, have excitation maxima well within the range of a Ti:sapphire laser. Their emission is shifted well into the visible spectrum (470-560 nm), making them ideal candidates for two-photon excitation.

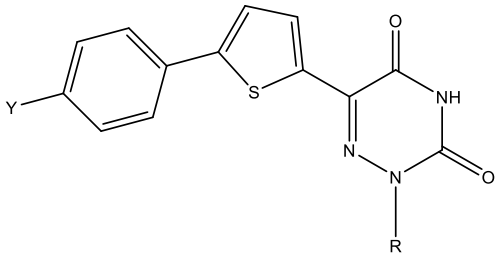


Figure 7-16. Extended thiophene-6-aza-uridines: 1a Y=F, 1b Y=H, 1c Y=Me, 1d Y=OMe, 1e Y=OH, 1f Y= NMe2. R=ribose.

The extended-6-azauridines quantum yields are very solvent-dependent and it is not yet known how they are affected by incorporation into oligonucleotides. Dioxane was chosen as the solvent for the present measurements, as all members of the extended-6-aza-uridine family have high quantum yields in it. In water, the quantum yield of 1e and 1f is < 0.01, which would lead to detection problems.

The steady-state fluorescence properties of the extended 6-aza-uridines in dioxane were studied following excitation at a laser wavelength anticipated to allow two-photon excitation. To confirm that absorption was occurring *via* a two-photon process, fluorescence intensity was measured as a function of laser power. The resulting log-log plot for 1f in dioxane, following excitation at 840 nm is shown in Figure 7-17. A slope of 1.85 ± 0.04 confirms a two-photon process.

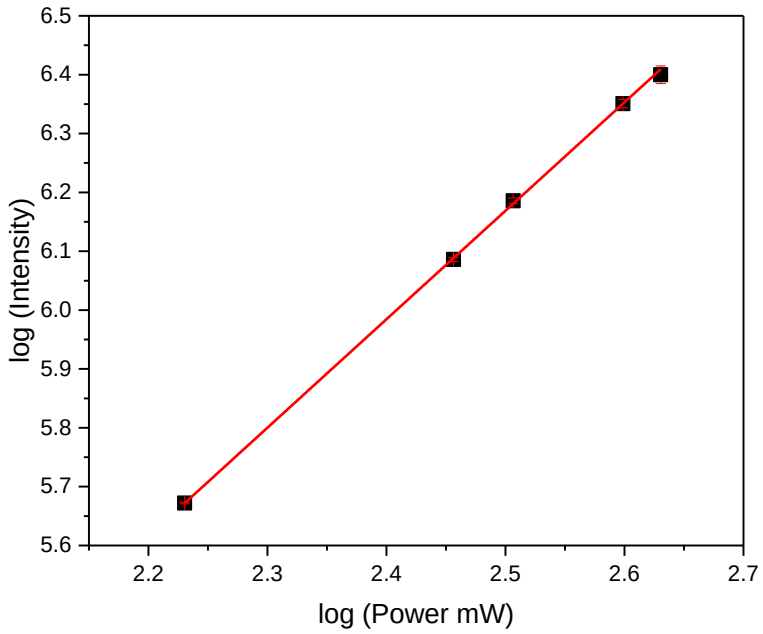


Figure 7-17. Log-log plot of fluorescence intensity vs laser power for 1F in phosphate buffer, Excitation wavelength was 840 nm. Laser power was varied from 170 mW to 427 mW.

Figure 7-18 shows the emission spectrum for each of the extended 6-aza-uridines 1a, 1c, 1d, 1e and 1f, resulting from two-photon excitation and the corresponding emission spectrum resulting from one-photon excitation. Two-photon excitation results in the same emission spectra as those seen for one-photon excitation. This indicates that unlike pA, emission is dominated by the same species in each instance.

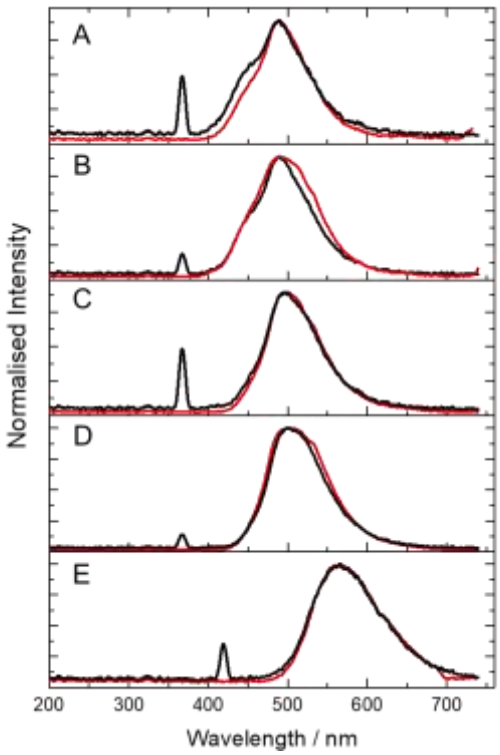


Figure 7-18. Emission spectra of the extended 6-aza-uridines in dioxane, following one-photon and two-photon excitation. A) 1a; excitation 750 nm (red), excitation 370 nm (black). B) 1c; excitation 780 nm (red), excitation 370 nm (black). C) 1d; excitation 780 nm (red), excitation 370 nm (black). D) 1e; excitation 780 nm (red), excitation 370 nm (black). E) 1f; excitation 850 nm (red), excitation 420 nm (black).

The cross-sections of the extended-6-aza-uridines in dioxane, calculated relative to rhodamine 6G, are reported in Table 7-11. The cross-section of rhodamine 6G, at this excitation wavelength, was confirmed by measuring it relative to coumarin 153 in DMSO.

	Substituent	σ_2/GM at (wavelength)	ϕ in **dioxane**, water	$\phi\sigma_2$ / GM in **dioxane**, water*
1a	F	5.1 ± 0.6 (750 nm)	**0.32**	**1.6**
			0.24	1.2
1c	Me	6.3 ± 0.4 (770 nm)	**0.38**	**2.4**
			0.12	0.75
1d	OMe	9.2 ± 0.4 (780 nm)	**0.74**	**6.8**
			0.02	0.18
1e	OH	14.2 ± 0.6 (780 nm)	**0.71**	**10.1**
			<0.01	<0.14
1f	NMe_2	90.0 ± 5 (840 nm)	**0.20**	**18.0**
			<0.01	<0.90

Table 7-11. Two-photon cross-section (σ_2) and two-photon brightness ($\phi\sigma_2$) of extended 6-aza-uridines 1a, 1c,1d,1e and 1f in dioxane, measured relative to rhodamine 6G in methanol. Error in the standard values of σ_2 is reported as 8%. Error in cross-section calculated from the standard deviation of measurements made at differing power. Values in bold indicate those measured in dioxane, those not in bold indicate values in water.* Values in water calculated assuming that the cross-section is not solvent-dependent.

The two-photon cross-sections of the extended-6-aza-uridines are interesting for two reasons. Firstly, the cross-sections measured are higher than those of any other FBAs studied previously, including pA. In dioxane the two-photon brightness values of 1d, 1e and 1f in dioxane are all higher than the brightness of pA (in ethanol). 1f has a brightness of 18 GM which is three times greater and a cross-section in dioxane of 90 GM, an order of magnitude higher than that of pA. Secondly, the cross-section increases with increasing electron-donating ability of the functional group (Figure 7-16). By creating a "push-pull" effect, the two-photon cross-sections have been increased, as has been achieved previously in larger dye molecules. The addition of an electron-donating group, even on a small structure can act to increase the cross-section by an order of magnitude. There have been no previous measurements of two-photon cross-sections of fluorescent base analogues with systematic variation of functional groups. By measuring the effect of the substituent on cross-section, a strategy for improving 2P cross-section in FBAs by the addition of electron-withdrawing groups, can be created.

In water, assuming that the two-photon cross-section is constant and only the quantum yield changes, the brightness of all the extended-6-aza-uridines would be lower than that observed for pA. Further measurement of two-photon cross-sections in oligonucleotides are needed to reveal whether the two-photon brightness of extended-6-aza-uridines in oligonucleotides would be more similar to those in dioxane or water. The quantum yield is expected to be closer to that of dioxane than water, due to the similar polarity but there is also the effect of the electrostatic environment to consider. As inclusion of pA in an oligo resulted in a decrease in the 2P cross-section it is possible a similar effect would also be seen for the extended-6-aza-uridines.

7.4 Conclusions and Future Work

The fluorescent base analogues discussed here have the highest two-photon cross-sections and brightness values of any FBAs previously reported.

Pentacyclic adenine has been shown to have the highest two-photon brightness reported to date for a FBA in an oligonucleotide. To our knowledge, there is only one previous report of the measurement of the two-photon cross-section of a FBA in an oligonucleotide. This was for tC, an environmentally insensitive FBA, with a two-photon brightness of 0.32 GM in a single-stranded oligonucleotide. Although the two-photon cross-section of pA in oligonucleotides is lower than that of the free base, depending on sequence context it is still as bright as many currently available FBAs free in solution. The cross-section values are similar for single and double strands and do not show any significant dependence on sequence context. The two-photon brightness shows much greater variation, with sequence context due to the sensitivity of the quantum yield to interaction with the neighbouring bases.

The brightness of pA, together with favourable excitation and emission wavelengths, gives the potential for use in two-photon microscopy and makes pA a promising candidate for single-molecule detection.

To elucidate whether the difference in one and two-photon spectral shape of pA is due to tautomerism, it would be advantageous to measure the fluorescence lifetimes under two-photon excitation. If there were similar lifetimes for two-photon excitation at 780 nm and one-photon excitation at 390 nm, but with different A-factors, this hypothesis would be confirmed.

The lower two-photon cross-section of pA in oligonucleotides has been attributed to the effect of the local electrostatic field in DNA. This effect could be explored to a greater extent. Goodson *et al.* only investigated the effect of intercalation into DNA using one dye.[31] A study into more intercalating dyes and into more fluorescent base analogues in DNA could increase the understanding of this effect on the two-photon cross-section.

The qAN compounds were found to be less favourable for two-photon excitation than pA. Although the maximum two-photon cross-section could not be measured, due to the wavelength limitation of the experimental set-up, it does not appear their cross-sections will be higher than pA, and their lower quantum yields makes them less desirable for two-photon use.

The extended-6-aza-uridines on the other hand show great potential for use in two-photon spectroscopy. A two-photon cross-section of 90 ± 5 GM for 1f is a significant improvement on any other FBA. As these FBAs have never been studied in oligonucleotides it is impossible to predict how the two-photon brightness will be affected by inclusion. Should the quantum yield remain high, as it is in dioxane, this could be a major advance. Measuring the quantum yield and two-photon cross-section of 1f in an oligonucleotide is an important next step.

The functional group found at a remote but conjugated position in the extended-6-aza-uridines, has been shown to have a dramatic effect on the cross-section. Similar modifications could be implemented to systematically improve the two-photon properties of other FBAs.

7.5 References

1 M. Göppert-Mayer, PhD Thesis University of Göttingen, 1931.

2 B. R. Masters and P. T. C. So, *Handbook of Biomedical Nonlinear Optical Microscopy*, Oxford University Press, New York, United States, 1st edn., 2008.

3 W. Denk, J. H. Strickler and W. W. Webb, *Science*, 1990, **248**, 73–79.

4 K. Svoboda and R. Yasuda, *Neuron*, 2006, **50**, 823–839.

5 M. Oheim, D. J. Michael, M. Geisbauer, D. Madsen and R. H. Chow, *Elsevier*, 2006, **58**, 788–808.

6 R. K. P. Piston and D. W. Benninger, *Curr Protoc Cell Biol.*, 2013, 1–36.

7 R. Weissleder, *Nat. Biotechnol.*, 2001, **19**, 316–317.

8 J. R. Lakowicz, *Principles of Fluorescence Spectroscopy Principles of Fluorescence Spectroscopy*, Springer, 3rd edn., 2006.

9 S. L. Jacques, *Phys. Medcine Biol.*, 2013, **58**, 37–61.

10 V. E. Centonze and J. G. White, *Biophys. J.*, 1998, **75**, 2015–2024.

11 K. König, in *Handbook of Biological Confocal Microscopy*, ed. J. Pawley, 3rd edn., 2006, pp. 680–689.

12 R. S. K. Lane and S. W. Magennis, *RSC Adv.*, 2012, **2**, 11397.

13 R. S. K. Lane, R. Jones, R. W. Sinkeldam, Y. Tor and S. W. Magennis, *Chemphyschem*, 2014, **15**, 867–71.

14 R. J. Stanley, Z. Hou, A. Yang, M. E. Hawkins, T. U. V, B. Hall and V. Pennsyl, *J. Phys. Chem. B*, 2005, **7**, 3690–3695.

15 A. Mikhaylov, J. R. Lindquist, P. R. Callis, B. Kohler, J. Pahapill, S. de Reguardatid, M. Rammod, M. Uudsemaad, A. Trummald and A. Rebane, *SPIE Proc.*, 2017, **10069**, 1–11.

16 A. C. Jones and R. K. Neely, *Q. Rev. Biophys.*, 2015, **2**, 1–36.

17 S. De Reguardati, J. Pahapill, A. Mikhailov, Y. Stepanenko and A. Rebane, *Opt. Express*, 2016, **24**, 9053–9066.

18 M. E. Hawkins, *Fluorescent pteridine probes for nucleic acid analysis.*, Elsevier Inc., 1st edn., 2008, vol. 450.

19 P. Sandin, L. M. Wilhelmsson, P. Lincoln, V. E. C. Powers, T. Brown and B. Albinsson, *Nucleic Acids Res.*, 2005, **33**, 5019–25.

20 C. C. Goh, L. G. Ng, J. Zhou, Z. Zhao, B. Liu, B. Z. Tang, J. Xiang, X. Cai, X. Lou, G. Feng, X. Min, W. Luo and B. He, *Appl. Mater. Interfaces*, 2015, **7**, 14965–14974.

21 M. Pawlicki, H. A. Collins, R. G. Denning and H. L. Anderson, *Angew. Chemie Int. Ed.*, 2009, **48**, 3244–3266.

22 M. S. Rocha and L. A. Reis, *Biopolymers*, 2017, **107**, 1–8.

23 Cornell University Two-Photon Cross Sections Database, http://www.drbio.cornell.edu/cross_sections.html, (accessed August 2017).

24 W. R. Zipfel, R. M. Williams and W. W. Webb, *Nat. Biotechnol.*, 2003, **21**, 1369–1377.

25 University of South Florida Microscopy, http://health.usf.edu/medicine/corefacilities/microscopy/olympus-fv1000 (accessed August 2017).

26 W. A. Franklin and J. D. Locker, *J. Histochem. Cytochem.*, 1981, **29**, 572–576.

27 H. Malak, F. N. Castellano, I. Gryczynski and J. R. Lakowicz, *Biophys. Chem.*, 1997, **67**, 35–41.

28 F. Bestvater, E. Spiess, G. Stobrawa, M. Hacker, T. Feurer, T. Porwol, U. Berchner-Pfannschmidt, C. Wotzlaw and H. Acker, *J. Microsc.*, 2002, **208**, 108–115.

29 M. Drobizhev, S. Tillo, N. S. Makarov, T. E. Hughes and A. Rebane, *J. Phys. Chem. B*, 2009, **113**, 12860–12864.

30 M. Drobizhev, N. S. Makarov, S. E. Tillo, T. E. Hughes and A. Rebane, *Nat. Methods*, 2011, **8**, 393–399.

31 P. H. Doan, D. R. G. Pitter, A. Kocher, J. N. Wilson and T. Goodson, *J. Am. Chem. Soc.*, 2015, **137**, 9198−9201.

32 D. Magde, R. Wong and P. G. Seybold, *Photochem. Photobiol.*, 2007, **75**, 327–334.

33 A. M. Brouwer, *Pure Appl. Chem.*, 2011, **83**, 2213–2228.

34 S. Sakadži, U. Demirbas, T. R. Mempel, A. Moore, S. Ruvinskaya, D. A. Boas, A. Sennaroglu, F. X. Kartner and J. G. Fujimoto, *Opt. Express*, 2008, **16**, 3341–3349.

Chapter 8: Towards Single Molecule Detection of Fluorescent Base Analogues

8.1 Introduction

Detection of individual molecules represents the ultimate level of sensitivity in any measurement. Details about a molecular system that would be concealed using conventional ensemble techniques can be revealed using single-molecule methods.

While an ensemble technique such as TCSPC can give information on, for example, the average proportion of extra-helical bases in a certain DNA sequence, it cannot inform on whether some strands have an extra-helical base for the duration of the measurement and some remain fully stacked or whether many oligonucleotides are interconverting between conformations. A single molecule measurement would be able to inform on the length of time a specific base is extra-helical. For this reason, finding a fluorescent base analogue that is bright enough to detect at the single molecule level has been an area of interest for a number of years.

2AP was assessed using fluorescence correlation spectroscopy (FCS) by Rigler *et al.* in order to determine its feasibility for use in single-molecule studies.[1] A count rate of 2 kHz per molecule was found with a background intensity of 7 kHz resulting in a signal-to-background ratio of 0.3.[1] These measurements were made on the free 2AP base exciting at 300 nm and using a power of 0.4 mW. These results indicate 2AP would not be appropriate for single molecule use, especially considering the degree of quenching it undergoes in DNA which would act to reduce brightness even more. A second fluorescent base analogue 3-MI was investigated by Sanabia *et al.*[2] 3-MI appears to be a more promising candidate for single molecule investigations as it has a count rate of 4 kHz, which was approximately five times greater than the background value. 3-MI in a 36-mer oligonucleotide was investigated at the same time and a count rate of 2 kHz per molecule was found, just higher than the value of the background. Sanabia *et al.* suggested two-photon excitation as a possible method to increase the signal-to-background ratio.

A paper published in 2013 by Rueda *et al.* claims to show the first single molecule detection of 2AP and pyrrolocytosine in DNA *via* UV excitation on a surface.[3] For 2AP in single and double-stranded DNA, "blinking" was observed and this was attributed to 2AP base flipping. Due to the counts per molecule and a signal-to-background ratio for free-2AP of less than 1 reported by Rigler *et al.* it is highly surprising that this was detected, especially as excitation was at 325 nm, not optimum for 2AP. Even more surprising is detection of pyrrolocytosine a

FBA with a quantum yield of only 0.04. Should Rueda *et al.* have succeeded in lowering the background to a point where single molecules of 2AP in DNA could be detected this would be a significant finding. However, no results pertaining to controls were shown. Background contamination is known to be significantly problematic when working in the UV. If conclusive background measurements were included, these results would be more convincing. No other groups have reported similar results in the 5 years since publication.

Many of the advantages of two-photon excitation were previously discussed in Chapter 7. For the detection of single FBAs that require excitation in the UV, two-photon excitation can be advantageous as it leads to a lower background. Two-photon spectroscopy has therefore been used to characterise several FBAs.

A study into the two-photon excitation of 6-MI showed that while a FCS curve could be collected for free 6-MI, it was not possible to measure a FCS curve when 6-MI was in an oligonucleotide due to the low quantum yield.[4,5] A value for counts per molecule was not reported so it is not possible to know if a single molecule of 6-MI would be detectable. Lane and Magennis assessed the ability of the environmentally insensitive FBA, tC, to be used in microscopy.[6] They showed that for tC in an oligonucleotide a similar count rate could be achieved for both 1P and 2P excitation and that 2P excitation was more favourable due to lower levels of photobleaching. Unfortunately, signal was too low to obtain a correlation curve but a count rate of ~0.1 kHz per molecule was estimated, approximately the same value as that of the background.

With the improved 2P brightness of the molecules discussed in Chapter 7, it may be feasible to detect single molecule bursts using 2P excitation. Comparisons can be made to coumarin 120 (C120): a molecule that absorbs at 350 nm; has a two photon cross-section of 3 GM and a quantum yield of ~0.50. [5,7] Single molecules of C120 were detected in both water and triacetin, although the average CPM appeared to be only ca. 0.5 kHz.[7] Interestingly this study demonstrated that the signal-to-background ratio was higher for C120 when two-photon excitation was employed than when 1P was used. As pA in an oligonucleotide has a similar two-photon brightness to C120, this suggests that pA may be possible to detect, in DNA, at a single molecule level.

Measurements in this Chapter have been made using ultrafast, shaped laser pulses for excitation. Temporally shorter pulses allow high peak power while maintaining a low average power. Previously these ultrafast pulses, < 10 fs, have been shown to reduce photobleaching and give a higher signal-to-noise ratio.[8,9] Through use of pulse shaping, selective excitation designed to optimise the absorbance of a specific molecule can also be achieved.[10,11]

Ultrafast pulses have a broad spectrum, as spectral width is inversely proportional to pulse duration. Assuming a Gaussian pulse shape, for a pulse centred about 800 nm with a pulse duration of 10 fs, this results in a bandwidth of ~ 100 nm. As a pulse this broad passes through, for example, a microscope objective, it will suffer from group velocity dispersion. A multiphoton intrapulse interference phase scan (MIIPS) system has been used in this work to compensate for chromatic phase dispersion caused by the broadband pulses passing through the microscope and control pulse compression and shaping.[12] The basic principle of optical pulse shaping is that a femto second pulse is first decomposed, using a spectral disperser such as a grating and a lens or objective, into its constituent parts. The phase and amplitude of the spectral components are then modulated by a spatially patterned mask before being recombined by a second grating and lens. The pulse shape is determined by the Fourier transform of the pattern of the mask with the spectral phase. Masks are implemented using a liquid crystal spatial light modulator array. In order to correct for phase distortion caused by the microscope objective, the dispersion must first be measured. MIIPS measures the second harmonic generated response to a series of spectral phase functions being applied to the pulses. This response elucidates the dispersion of the pulse caused by the objective and a mask that cancels this distortion can be applied.

From the results of Chapter 7, pA was identified as having the highest two-photon brightness of any fluorescent base analogue previously measured in an oligonucleotide. Its two-photon brightness is similar to that of coumarin 120, a molecule that has a similar excitation wavelength and has been detected at the single molecule level. This suggests that detection of pA should also be possible. This Chapter presents the efforts made to detect pA in an oligonucleotide at the single molecule level using confocal microscopy. In Chapter 7 it was also shown that the extended-6-aza-uridine 1f has a two-photon cross-section of 90 GM, an order of magnitude greater than that of any other FBA. This chapter will therefore also include a preliminary assessment of the 1f nucleobase.

8.2 Experimental

8.2.1 Ensemble One-photon and Two-photon Photobleaching

A schematic of the set up used to compare one and two-photon photobleaching is shown in Figure 8-1. This is a modification of the set up used in Chapter 7.

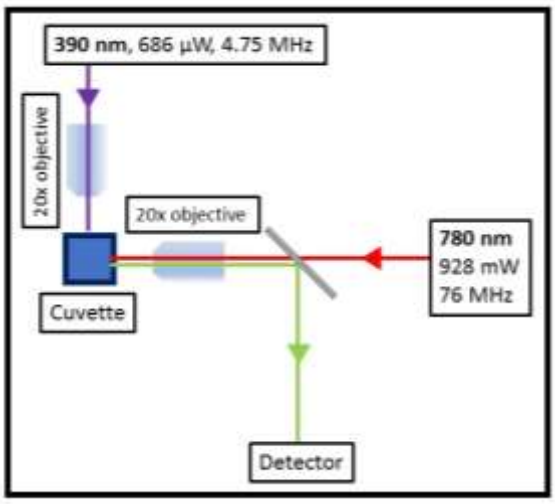


Figure 8-1. Schematic of set up used to compare one and two-photon photobleaching.

A mode-locked Ti:sapphire laser (Coherent Mira pumped by Coherent Verdi), producing pulses of duration ~ 150 fs at a repetition rate of 76 MHz, was used as the two-photon excitation source. The second harmonic, 390 nm, from this same laser was used for one-photon excitation. A pulse picker (Coherent 9200) was used to reduce the repetition rate to 4.75 MHz. Two-photon excitation has previously been described in Chapter 7. A variable reflective neutral density filter was used to attenuate the excitation beam, which then passed through a dichroic mirror (Semrock Brightline FF735-Di02) and was focused by a 20× objective (Olympus) into the sample solution, which was contained in a 1-cm pathlength cuvette. Fluorescence emission was collected by the same objective, reflected from the dichroic mirror, passed through a shortpass filter (Semrock Brightline FF01-720/SP-25) and detected by a fibre-coupled spectrometer (Ocean Optics USB2000+), with an acquisition time of 1 s. Measurements were recorded at 10 second intervals. One-photon excitation occurred at 90° to detection. The excitation beam was focused using a 20× objective (Olympus) into the sample with the excitation volume aligned so as to optimise collection from the second 20× objective (Olympus). Fluorescence was then reflected from the dichroic and collected in an identical manner to two-photon excitation. For both 1P and 2P excitation power was measured after the cuvette.

8.2.2 Two-Photon Microscopy

The set up used for 2P Microscopy is illustrated by the schematic shown in Figure 8-2.

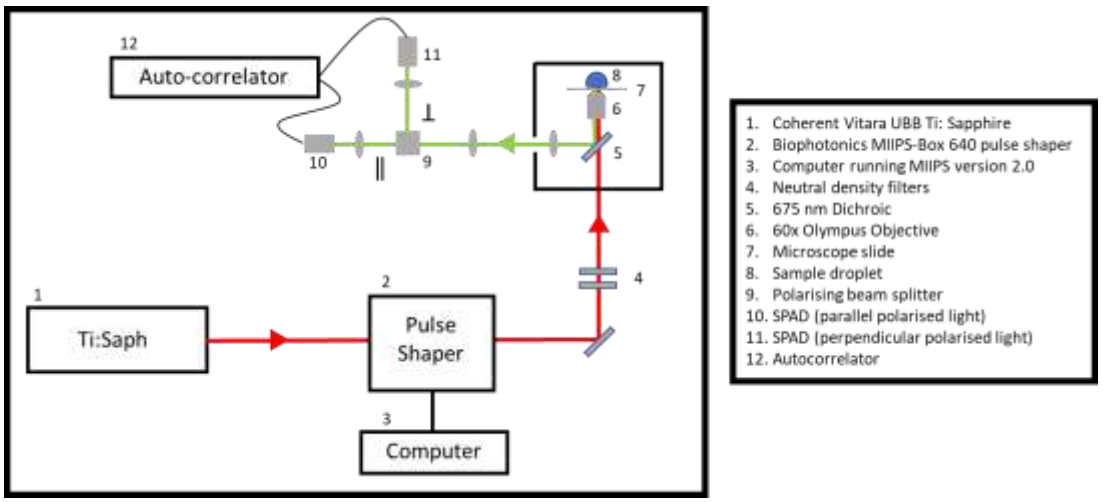


Figure 8-2.Schematic of fluorescence correlation spectroscopy set up.

A broadband Ti:Sapphire laser (Coherent, Vitara UBB) producing pulses (15 fs, FWHM of 135 nm centred at 800 nm) at a repetition rate of 80 MHz was used as the excitation source. This light is passed through a pulse shaper (Biophotonics, MIIPS-Box 640) containing two liquid-crystal spatial light modulators (Jenoptik, SLM-S640d). Pulse shaping is controlled by MIIPS software version 2.0 (Biophotonic Solutions). To determine the required pulse shape, phase must be measured. This is done by measuring the second harmonic spectrum using an Ocean Optics USB4000 spectrometer. To do this, during pulse shaping a SHG crystal (10µm BBO, Biophotonic Solutions Inc.) is placed at the sample position. After the pulse shaper the laser beam passes through two reflective neutral density filters (Thorlabs, NDK01) that are adjusted to control power and enters an inverted Olympus IX71 microscope. Light is then focused into a sample droplet (or the SHG crystal during pulse shaping) using a water immersive objective lens (Olympus, UPLAPO 60x). The solution samples were contained in chambered cover glasses (Nunc Lab-Tek II) and the solution temperature, kept at 21 ± 1 °C, was controlled with an incubator (Live Cell Instrument, CU-501). Emitted fluorescence is collected by the same objective used for excitation and it is transmitted by a NIR-reflecting dichroic (Chroma, 675 DCSPXR) towards the detector. Emitted photons were divided into parallel and perpendicular components by a polarising beam splitter before detection by two single photon avalanche photodiodes (MPD, PDM 50c and $PD-050-CTB) positioned perpendicular to one another. The signal from these detectors was collected by either a photon-counting card (SPC-132, Becker and Hickl GmbH) or a hardware correlator (ALV-7002, ALV GmbH).

For fluorescence correlation spectroscopy the correlation curve ($G(t)$) was fit, taking into account background contribution (I_B) and assuming diffusion through a 3D Gaussian volume using the equation:

$$G(t) = 1 + \frac{(1 - \frac{I_B}{S})^2}{\sqrt{8N}} \left(\frac{1}{1 + \left(\frac{t}{\tau_D}\right)} \right) \times \left(\frac{1}{1 + \left(\frac{\omega_0}{z_0}\right)^2 \left(\frac{t}{\tau_D}\right)} \right)^{1/2}$$

Equation 8-1

where N is the average number of molecules in the focal volume, S is total signal strength (background and sample), ω_0 and z_0 are the lateral and axial dimensions of the Gaussian volume and τ_D is the average diffusion time.

8.2.3 Sample Preparation for FCS

pA-containing oligonucleotides, 5'-d(CGCAG(pA)ATCG)-3', were diluted in 20 mM Tris (Sigma-Aldrich), 150 mM NaCl (Fluka), and 1mM $MgCl_2$ (Fluka), pH 7.2. Ultrapure water was used in the preparation of buffers. Buffers were filtered through activated charcoal to remove any fluorescent impurities. Measurements were made in chambered cover glasses. These were cleaned with methanol followed by ultrapure water prior to use. Background measurements of buffer were made prior to each sample measurement.

8.3 Results

8.3.1 Fluorescence Correlation Spectroscopy of pA

The first FBA studied with the aim of single molecule detection was pA. A crucial property for molecules detected at the single molecule level is photostability. Photobleaching is a common issue with fluorescent base analogues. During one-photon measurements of pA some degree of photobleaching had been observed. For 1P measurements photobleaching could be avoided by use of low powers. To determine whether photobleaching would also be an issue when the higher power required for two-photon excitation was used, a photobleaching study was conducted. To compare the stability of pA during 1P and 2P excitation, fluorescence emission from bulk sample was measured as a function of time, with excitation performed at either 780 nm (2P) or 390 nm (1P). Figure 8-3 shows typical bleaching curves for the pA riboside in water (10% EtOH). Power was adjusted such that initial emission signal intensity was the same each case. The 2P excitation power is three orders of magnitude higher than that required for 1P with a pulse rate 16 times greater. Over the timescale investigated 2P emission remains constant whereas 1P emission decayed exponentially. The 1P decay could be fit to a biexponential function with decay parameters of 67 s (58 %) and 870 s (42 %). pA appears to

be photostable under 2P excitation conditions which allowed single molecule detection to be attempted.

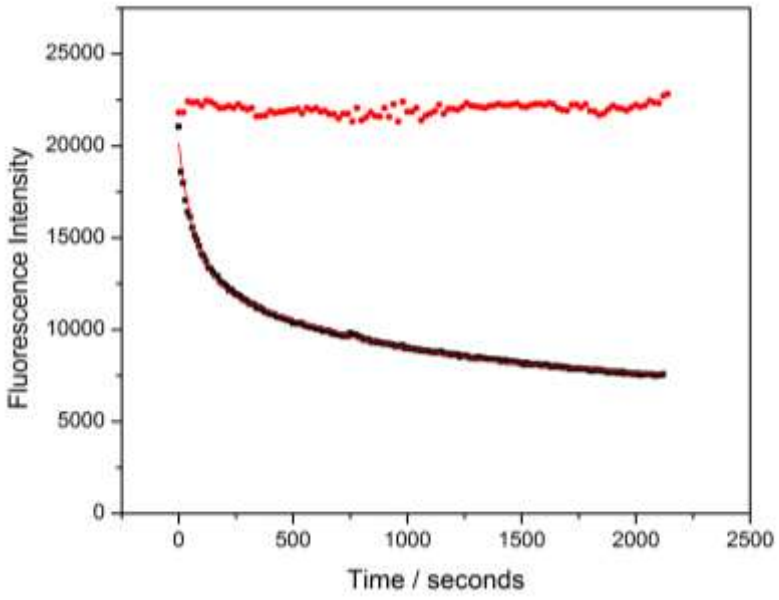


Figure 8-3. Comparison of the one-photon (black) and two-photon (red) emission of pA (H2O:EtOH, 90:10) as a function of time. 2P excitation at 780 nm from the fundamental of a Ti:sapphire, repetition rate 76 MHz, average power 928 mW. 1P excitation at 390 nm from the second harmonic of the same laser, repetition rate 4.75 MHz, average power 686 µW. Intensity measured at 420 nm.

From the series of 10-mer oligonucleotides investigated in Chapter 7 the sequence in which G and A are neighbouring bases to pA was identified as having the highest 2P brightness. This sequence was consequently chosen for use in all single molecule measurements, this pA containing single stranded oligonucleotide will be referred to as pA-ssGA.

The photostability of pA-ssGA, was further investigated under 2P microscopy conditions. Emission was split according to polarization and collected by two photon-counting detectors. The trace displayed in Figure 8-4 shows intensity over a 30 minutes period. The parallel channel relative to the laser is displayed in black and perpendicular in red. This trace corresponds to an average of four molecules in the focal volume and an excitation power equal to 11 mW. Intensity remains constant for the duration of the measurement indicating that photobleaching occurs on a timescale slower than diffusion of molecules through the focal volume.

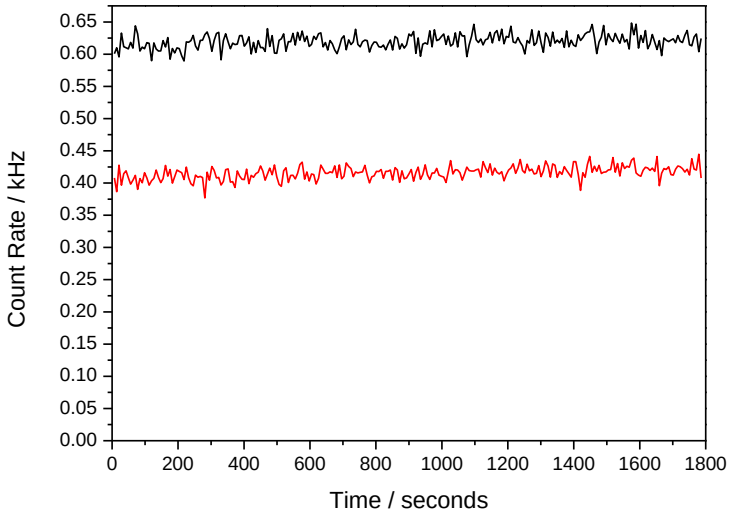


Figure 8-4. pA-GA oligonucleotide count rate over 1800 seconds. The emission signal is split according to polarization with respect to the laser; parallel in black and perpendicular in red.

By adjusting the spectral phase of the exciting pulse it is possible to selectively excite certain fluorophores. By adjusting the pulse shape it was hoped that it would be possible to increase signal from pA-ssGA without increasing signal from the background. To identify the optimum pulse shape for single molecule detection, a series of pulses were investigated. The full spectrum of the laser, 650 to 950 nm, is displayed in Figure 8-5. Spatial masking is controlled automatically through the MIIPS software and acts to block different parts of the spectrum. The blocked parts of the spectrum are represented in Figure 8-5 by the grey sections.

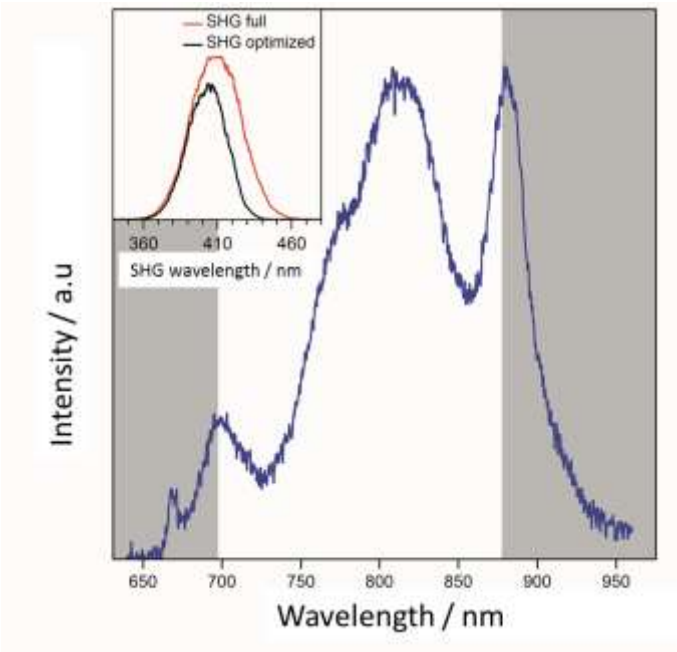


Figure 8-5. Spectrum of laser (blue). Grey areas indicate spectral components blocked for optimum pulse shape. The second harmonic generated spectrum for the full laser spectrum and the masked pulse are also displayed. Figure provided by David Nobis, University of Glasgow.

From these measurements the pulse shape leading to the highest signal-to-background ratio was identified. The highest signal-to-noise ratio achieved for each shape is presented in Table 8-1. The pulse shape is described in terms of the centre of the amplitude mask and the width. As power also affects signal-to-noise, and this varied when different masks were applied, each pulse shape was investigated across a range of powers. The values in Table 8-1 are those recorded using the optimum power, the power at which the SBR was at its highest.

Pulse Shape centre/width / nm	Power / mW	pA-ssGA Count Rate / kHz	Background Count Rate / kHz	Signal/Background
788/ 119*	11.3	7.5	1.1	5.8
757/ 119	11.4	6.2	0.76	7.2
725/ 121	11.5	5.7	1.1	4.2
819/ 118	11.0	4.2	0.67	5.3
788/ 87	11.0	6.5	0.74	7.8
788/ 150	11.0	7.7	0.67	10.5
787/ 181	11.9	8.1	0.61	10.7

Table 8-1. Count rate for pA, count rate of background and signal/background ratio for pA-GA when excited using pulses of different central wavelength and spectral range. * Optimum pulse shape with no spectral masking.

An amplitude mask centred at 787 nm with a width of 182 nm was identified as producing the maximum signal-to-noise ratio. The centre of this mask is consistent with the wavelength identified from bulk 2P measurements (Chapter 7) as the wavelength at which a maximum 2P cross-section was achieved. The second harmonic generated spectrum (Figure 8-5) from the shaped pulse is slightly blue shifted, closer to the maximum 1P absorption wavelength, compared to the unshaped pulse.

Compared to a compressed pulse with no additional shaping, the signal-to-noise ratio increases by almost 50% when a shaped pulse is used. SBR is 5.8 for the compressed pulse and 10.7 for the shaped pulse. The pA count rate only slightly increases when using the shaped pulse: 8.1 kHz compared to 7.5 kHz. It is possible this difference could be due in part to the difference in power, as measurements with the compressed pulse were made at 11.3 mW and with the shaped pulse at 11.9 mW. The main improvement when using a shaped pulse is background suppression. The background rate for the compressed pulse is 1.1 kHz and this is reduced to 0.61.

Using the pulse shape established from the results of Table 8-1, a power series was collected and used to identify the optimum power at which to make measurements. This power series is displayed in Figure 8-6.

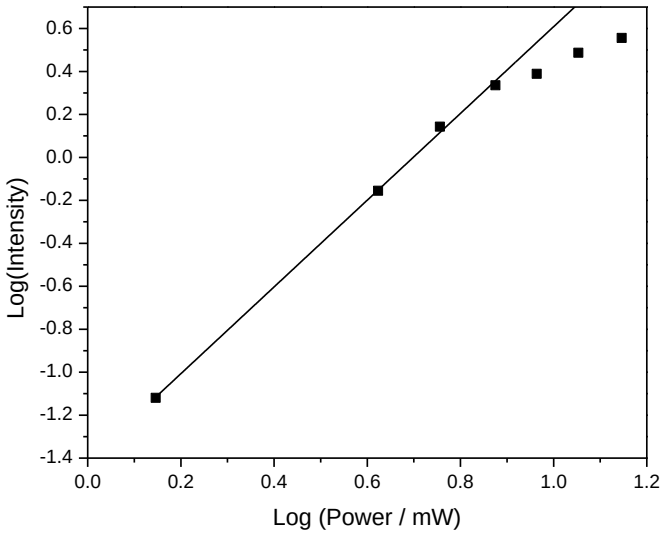


Figure 8-6. Logarithmic power dependence of pA- GA oligonucleotide with no additional pulse shaping. Slope of 2.02 on fitting first four point. Loss of linearity at powers greater than 7.5 mW.

For powers up to 7.5 mW a linear relationship, slope equal to 2.02, between log(Intensity) and log(Power) is observed confirming a two-photon process. The power at which the signal to background intensity was found to be highest was 11.9 mW. This power was used for all further measurements. Loss of linearity at high powers most likely occurs due to molecules already being in the excited state. At some saturation power, an increase in power will not result in further excitation and signal will plateau.

A representative FCS curve recorded for pA-ssGA, at 11.9 mW using the optimum pulse shape, is shown in Figure 8-7A. For comparison, an identical measurement was made using buffer. This is shown in Figure 8-7B.

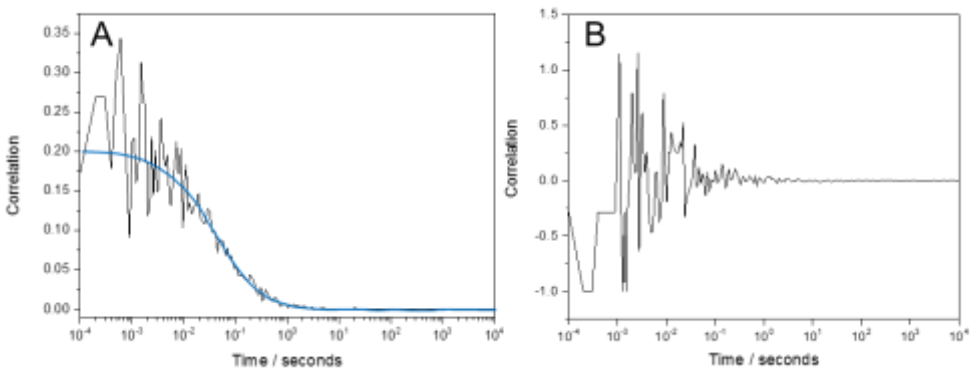


Figure 8-7. A) FCS pA-oligonucleotide (black), fit (blue), 11.5 mW 30-minute collection time. B) Background. Buffer run under same conditions, for same timescale.

From Figure 8-7B it can be seen that when excited at 11.5 mW for 30 minutes the buffer sample displays no correlation, indicating that no emitting contaminants are present.

From FCS curves such as that shown in Figure 8-7A the number of molecules in the focal volume was established. From this a value of counts per molecule could be calculated. To increase accuracy this was done for a series of pA concentrations. For pA-ssGA, this value was found to be 0.5 ± 0.03 kHz per molecule very slightly above the background of 400 Hz.

With the count rate of 0.5 ± 0.03 kHz and a background of 0.4 kHz the likelihood of detection of single molecule bursts is low. When coumarin 120 was investigated by Seidel *et al.* the count rate in that case was approximately 500 Hz/molecule, yet detection was possible due to a difference between average count rate and the peak count rate.[7] The peak count rate was ~ 100 times greater than the average count rate and consequently single molecule bursts were detected. To investigate the possibility of pA populations with different brightness multichannel scaler (MCS) traces were collected. The MCS trace for ~10 pM solution of pA-ssGA in buffer and buffer alone are displayed in Figure 8-8.

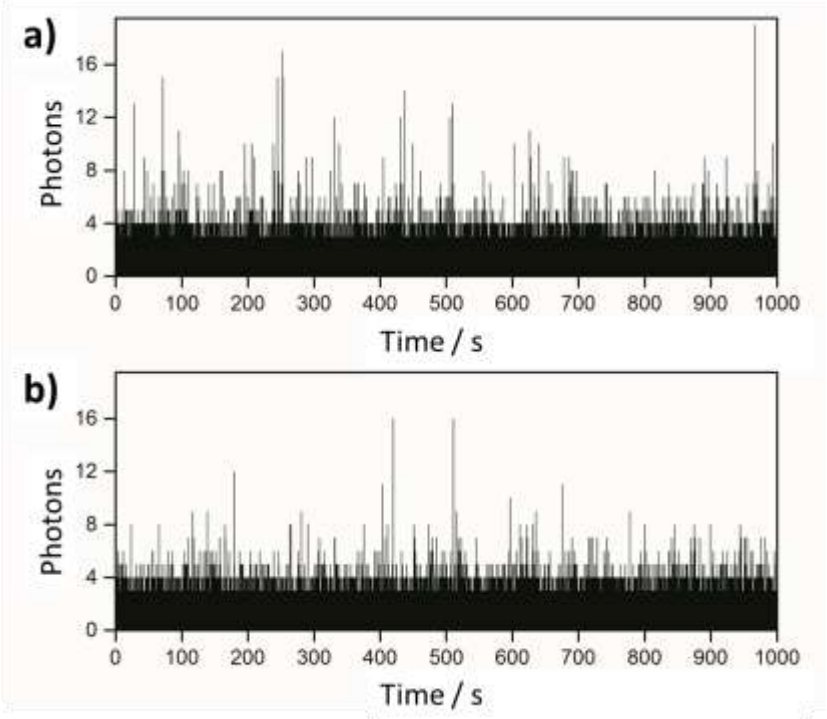


Figure 8-8. Confocal MCS trace for (a) pA-modified single-stranded oligonucleotide in buffer (1.1×10^{-11} M) and (b) pure buffer. Trace recorded at a power of 10.7 mW for 1000 seconds.

The MCS trace for the pA oligo and that recorded for pure buffer are not dramatically different. The density of bursts is greater for the sample than for buffer alone with slightly more high-

intensity bursts. A burst frequency histogram from the results shown in Figure 8-8 is displayed in Figure 8-9.

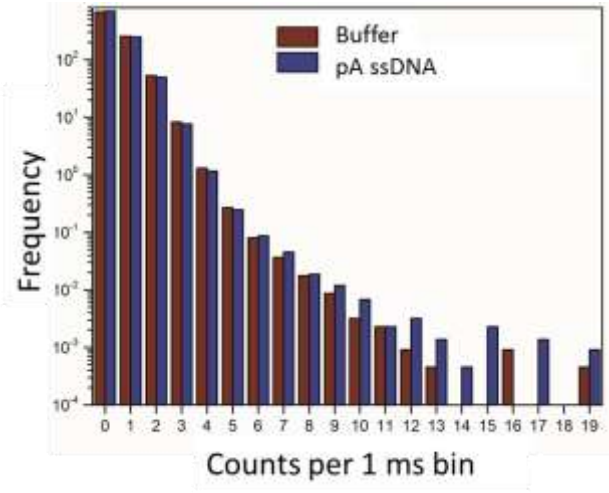


Figure 8-9. Photon counting histogram. Measurement length, 30 minutes. Laser power, 10.7 mW. Figure provided by David Nobis, University of Glasgow.

From the histogram it can be seen that the distribution and size of the bursts are similar to those in pure buffer. This shows that there are no sub-populations of pA states with different brightness and the histogram is consistent with an average count per molecule of 0.50 kHz.

8.3.2 Fluorescence Correlation Spectroscopy of 1f

Preliminary measurements have been made on the extended-6-aza-uridine 1f. In Chapter 7 it was shown that 1f has a cross-section of 90 GM in dioxane. Currently the only ultrapure solvent available to us for these measurements is water. In water the quantum yield is very low < 0.01 but nonetheless FCS measurements were made.

1f proved sufficiently stable and bright so an FCS curve could be collected. This FCS curve and a corresponding background measurement are displayed in Figure 8-10.

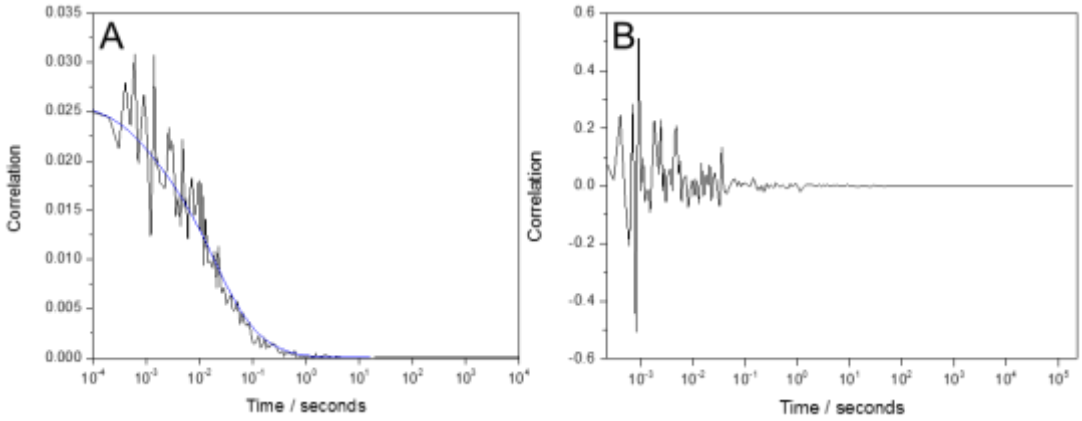


Figure 8-10. A) FCS curve of 1f in water (black), fit (blue), 19 mW, 2 minute collection time. B) Water background, collection time 20 minutes.

With approximately 40 1f molecules in the focal volume, within 120 seconds, a FCS of good enough quality to fit could be recorded. This is significantly shorter than the time-scale over which equivalent pA measurements were recorded. From the fit of the FCS curve shown in Figure 8-10 a count rate of 1.7 kHz per molecule of 1f was calculated with a background of 0.7 kHz. This value is already over three times greater than the counts per molecule found for pA. With optimisation of parameters such as pulse shape and power this value could be expected increase.

8.4 Conclusions and Future Work

For the first time, signal has been high enough and the molecule of interest photostable such that 2P FCS measurements of a fluorescent base analogue in an oligonucleotide could be made. These measurements were made using pentacyclic adenine in a 10-mer oligonucleotide with a sequence 5'-d(CGCAG(pA)ATCG)-3'. A count rate of 0.5 ± 0.03 kHz per molecule of pA was established at 11.4 mW. This value is of a similar magnitude to the background. These are highly promising results and there is much future work to be carried out.

There are several methods available to potentially increase single molecule brightness. These include the use of antifades such as Trolox and Vitamin C, both of which act to quench singlet oxygen. So far measurements made using antifades have resulted in correlation from the background. When vitamin C was used, an increase in brightness was found that was greater than that which could be attributed to just the vitamin C alone. This is highly promising and further attempts to purify vitamin C samples should be made.

Due to the pA nucleoside's insolubility in water and its propensity to aggregate, reliable FCS results for the pA riboside could not be obtained. For this same reason, single molecule bursts attributed to pA riboside in water could not reliably be recorded. pA is highly soluble in several other solvents, so measurements made in such solvents would be informative. This would allow a comparison between the count rate per molecule of free pA and pA in an oligonucleotide.

Measurements were also made on 1f. These measurements are in the very early stages. The count rate per molecule in water was found to be 1.7 kHz with a background of 0.79 kHz. A series of experiments to identify optimum pulse shape and power should be carried out. This is likely to lead to a higher signal-to-noise ratio, as was seen for pA.

For 1f in water the quantum yield is less than 0.01, in dioxane this value is significantly higher. As the polarity of dioxane is similar to that of the internal DNA environment, measurements

made in dioxane could give a more accurate indication of the count rate of 1f in DNA. It was shown in Chapter 7 that the pA 2P cross-section decreased in an oligonucleotide. It is possible a similar effect would be seen for 1f. Ultimately, measurements should be made on a 1f-containing oligonucleotide. pA was shown to be just below the brightness required to detect a single-molecule but this may prove to be possible with 1f.

8.5 References

1 S. Wennmalm, H. Blom, L. Wallerman and R. Rigler, *Biol. Chem.*, 2001, **382**, 393–397.

2 J. E. Sanabia, L. S. Goldner, P. A. Lacaze and M. E. Hawkins, *J. Phys. Chem. B*, 2004, **108**, 15293–15300.

3 E. A. Alemán, C. De Silva, E. M. Patrick, K. Musier-Forsyth and D. Rueda, *J. Phys. Chem. Lett.*, 2014, **5**, 777–781.

4 E. Katilius and N. W. Woodbury, *J. Biomed. Opt.*, 2006, **11**, 44004.

5 G. V. Prakash, C. Santhamma, S. Bangarraju and D. L. Sastry, *Indian J. Pure Appl. Phys.*, 2000, **38**, 635–644.

6 R. S. K. Lane and S. W. Magennis, *RSC Adv.*, 2012, **2**, 11397.

7 L. Brand, C. Eggeling, C. Zander, K. H. Drexhage and C. A. M. Seidel, *J. Phys. Chem. A*, 1997, **101**, 4313–4321.

8 C. Wang and A. T. Yeh, *J. Biomed. Opt.*, 2012, **17**, 25003.

9 P. Xi, Y. Andegeko, L. R. Weisel, V. V. Lozovoy and M. Dantus, *Opt. Commun.*, 2008, **281**, 1841–1849.

10 E. R. Tkaczyk, A. H. Tkaczyk, K. Mauring, J. Y. Ye, J. R. Baker and T. B. Norris, *J. Fluoresc.*, 2009, **19**, 517–532.

11 M. H. Brenner, D. Cai, J. A. Swanson and J. P. Ogilvie, *Opt. Express*, 2013, **21**, 17256.

12 V. V. Lozovoy, I. Pastirk and M. Dantus, *Opt. Lett.*, 2004, **29**, 775.

Chapter 9 : Conclusions

The main aim of this thesis has been to characterise novel fluorescent base analogues, both their one- and two-photon properties, in order to establish if any prove to be more advantageous than those currently in use and to ascertain if any could be suitable for use in single-molecule studies.

Chapter 4 focused on the extended-6-azaurdines. These molecules were found to exhibit multiple fluorescence lifetimes in polar and non-polar solvents and these lifetimes were rationalised with the use of DFT. Two lifetimes were attributed to rotational isomers and it was hypothesised that a further lifetime, present in certain cases, could be the result of a tautomer. In Chapter 5 it was the quadracyclic adenines (qANs) that were investigated, lifetime measurements also indicated multiple emitting species, in this case thought to be tautomers. Excited-state proton transfer was identified as a mechanism by which tautomerism occurs for the qANs. The final fluorescent base analogue investigated in this thesis was pentacyclic adenine (pA). Due the promisingly high quantum yield exhibited by pA across a range of solvents, pA was also investigated in a series of oligonucleotides. The multiple lifetimes observed for pA in the oligonucleotides can most likely be assigned to multiple stacked conformations that pA adopts in relation to its neighbouring bases.

In the introduction of this thesis the ideal FBA was described as a molecule that can: be incorporated in DNA with no structural perturbation; has a high brightness; has good photostability and has absorption and emission in the visible region. Overall, it can be concluded than none of the FBAs investigated in this thesis fulfil all the required criteria of an ideal FBA, but several show improvements on those currently commercially available. The extended-6-azauridine, qAN1, qAN4 and pA have absorption spectra in a region that makes them accessible to two-photon excitation using a Ti:Sapphire laser. Both pA and all the extended-6-azauridine, most especially 1f, have been shown to have a 2P brightness higher than any FBA previously investigated. The two FBAs, pA and 1f, studied so far have also proven to have high photostability under the excitation condition required for 2P microscopy.

The most noticeable difference between the FBAs studied in this thesis and ones such as 2AP is their size. None of those studied could be described as truly isomorphic. Although these larger FBAs have, where investigated, been shown to hydrogen bond with complementary bases and not cause gross distortion to the helix, it has not been shown if dynamics are affected.

The increased size of these FBAs means that, most likely, they will not behave in the same manner as a natural base in terms of the conformations adopted, the nature of their interactions with neighbouring bases or interactions with other molecules, such as enzymes. For example, it is highly unlikely that a FBA such as pA would be base-flipped by a methyl transferase enzyme, whereas 2AP is able to mimic a natural base in this respect. One of the main advantages of a FBA over an externally linked dye is the ability to report on the internal structure of DNA. However, if the probe becomes big enough to distort the DNA structure, but without the high brightness of a dye, such a rhodamine, then this advantage is lost. FBAs with larger structures have the advantage of red-shifted absorption and higher quantum yield, but with these improvements in wavelength and brightness there is the inevitable trade-off of reduced isomorphism.

Ideally, a probe will have a single fluorescence lifetime when free in solution. All those studied in this thesis have been shown to exhibit multi-exponential decays. As was demonstrated for pA in Chapter 6, this makes analysis of the decay in an oligonucleotide complex. Unlike 2AP, where the different lifetime components seen in a strand of DNA can be attributed in a consistent manner to different stacking environments, some lifetimes exhibited by pA could be due to the presence of an intrinsic isomer of the base itself. It is unlikely that a FBA like pA, the qANs or the extended 6-aza-uridines will be of use in the same manner as 2AP, the lifetime distributions of which give insight into the conformational properties of DNA. The quantum yield of pA is, however, highly sensitive to the identity of its immediate neighbouring bases, it is therefore likely that it will be sensitive to base-flipping of a neighbouring base. As the quantum yield of pA in an oligonucleotide is higher than that of 2AP, it may be possible to quantify the intensity response to base flipping using steady-state methods. This could make it more widely applicable for kinetic studies of base flipping (and other enzyme-induced distortions of DNA) than 2AP, and investigation into this is an area of potential future work.

The use of DFT to rationalise the origin of multiple lifetimes has been shown to be useful. Rotational isomers were shown to exist for the 6-aza-uridines in Chapter 4, and excited state tautomers for the qANs in Chapter 5. As some level of TD-DFT-assisted rational design was already implemented as part of a screening process to select which qAN-based FBAs to synthesise, calculation of the relative energies of likely tautomers could also be implemented prior to synthesis, to help identify molecules likely exist in multiple forms. This thesis has shown that simple DFT calculations, such as these, could be used advantageously to inform the synthesis of future FBAs.

In terms of single-molecule detection of a FBA in DNA, the present studies on pA show that this should be possible. pA, with a two-photon brightness in an oligonucleotide of just less than 3 GM falls just below the limit of detection. Its emission wavelength is below 450 nm; should this be red-shifted further into the visible, where background is lower, or the cross-section increased slightly, detection of a single molecule should become possible. With a cross-section of 90 GM, 1f is an even more promising candidate than pA for single-molecule detection, and with some optimisation of excitation parameters could prove to be suitable. The extended-6-aza-uridines also exemplify how changing just one substituent can have a dramatic effect on two-photon cross-section and could act as a model for future FBA design. It can now be believed that, with concerted synthetic effort, FBAs with 2P brightness high enough for single-molecule detection could be produced and that single-molecule detection of FBA-containing DNA has potential to become a routine procedure. In summary, two-photon excitation has been shown to be a potentially valuable tool for single-molecule detection and increasing the two-photon cross-section of FBAs, possibly through the addition of electron-donating substituents, is a desirable and feasible objective.

Appendix I: Functional Choice 6-aza-uridine Parent

To ascertain the best functional and basis set to use benchmarking calculations were carried out. This was done in two ways: comparison of bond lengths to the crystal structure and comparison of the relative energy of conformer 1 and 2 to those calculated using MP2.

Four bond lengths, calculated using various functional/basis set combinations, have been included in Table 1 along with the equivalent bond lengths from the crystal structure.

$$\text{Total Bond Length Error} = \frac{1}{n}\sum_{i=1}^{n}\left|x_{crystal} - x_{dft}\right|$$

Equation I-1

Where $x_{crystal}$ is the bond length from the crystal structure and x_{dft} is the DFT bond length.

Functional/ Basis Set	S1-C10	C12-O2	C13-N6	Total Bond Length Error	S1-O2 Distance	Dihedral angle S1-C10-C11-C12
Experimental	1.744	1.232	1.373	-	2.749	5.84
B3LYP/ 6-311g(d,p)	1.754	1.216	1.388	0.167	2.840	-0.001
B3LYP/ 6-311+g(d,p)	1.753	1.217	1.384	0.164	2.846	0.10
B3LYP/ 6-311++g(d,p)	1.753	1.217	1.384	0.164	2.846	0.10
B3LYP/SVP	1.751	1.216	1.393	0.191	2.829	-0.002
B3LYP/TVZP	1.759	1.216	1.383	0.161	2.848	0.18
M06-2X/ 6-311g(d,p)	1.735	1.208	1.380	0.160	2.834	0.03
M06-2X / 6-311+g(d,p)	1.735	1.209	1.378	0.159	2.846	0.003
M06-2X / 6-311++g(d,p)	1.735	1.209	1.378	0.158	2.842	0.01
wB97XD/ 6-311g(d,p)	1.737	1.210	1.379	0.152	2.847	-0.17
wB97XD / 6-311+g(d,p)	1.738	1.211	1.377	0.146	2.852	0.01
wB97XD / 6-311++g(d,p)	1.738	1.211	1.377	0.146	2.852	-0.10
BH&HLYP/ 6-311g(d,p)	1.734	1.200	1.370	0.169	2.839	0.001
BH&HLYP/ 6-311+g(d,p)	1.735	1.201	1.366	0.168	2.840	0.50

Table I-1. Bond lengths, total bond error, S-O distance and dihedral angle calculated for 6-aza-uridine parent using various functional/basis set combinations.

The bond lengths chosen were those between carbon and a hetero-atom specifically Carbon (10) to Sulphur (1), Carbon (12) to Oxygen (2) and Carbon (13) to Nitrogen (6). The Sulphur (1) to Oxygen (2) through space distance is also reported as it is illustrative of the orientation of the two aromatic rings. The dihedral angle between rings is included. Mean absolute error in bond length was calculated using Equation I-1 and the results are shown in Table 1. All

bond lengths were included in the error calculation except those involving a Hydrogen atom. All calculations were performed in a PCM dioxane solvent.

The total bond error, the S-O distance and dihedral angle are similar for all basis sets. This makes it hard to identify the best functional to use from this comparison alone. The crystal structure also contains a bromine atom whereas the compound used in spectroscopic studies contains a hydrogen. These factors lead to the conclusion that bond length is not the best metric by which to benchmark this compound.

As the property of interest from these calculations is the relative energy of the two conformers, this was instead used as a comparison. DFT results were compared to those calculated using a quantum mechanical approach (MP2) as this is known to give more reliable energies[11,12,13].

	Conformer 1 population	Conformer 2 population	Height of barrier /kJ mol^{-1}	Lifetime/ns for $C_1 \rightarrow C_2$
MP2/6-311g(d,p)	0.85	0.15	26.6	8
B3LYP/ 6-311g(d,p)	0.80	0.20	29.5	23
B3LYP/ 6-311+g(d,p)	0.82	0.18	28.5	16
B3LYP/ 6-311++g(d,p)	0.81	0.19	28.6	16
B3LYP/TVZP	0.79	0.21	28.4	16
CAM-B3LYP/ 6-311+g(d,p)	0.89	0.11	24.4	3
M06-2X/ 6-311g(d,p)	0.70	0.30	24.5	3
M06-2X / 6-311+g(d,p)	0.67	0.33	23.0	2
M06-2X / 6-311++g(d,p)	0.66	0.34	23.0	2
w/ 6-311g(d,p)	0.76	0.24	25.3	4
w / 6-311+g(d,p)	0.68	0.32	23.9	3
w / 6-311++g(d,p)	0.67	0.33	22.2	3
BH&HLYP / 6-311+g(d,p)	0.78	0.22	26.4	7

Table I-2. Population of conformer 1 and 2, barrier height between conformer 1 and 90 transition state and corresponding lifetime calculated using various functional/basis set combinations.

The choice of functional has a significant effect on the relative energy of the two populations and the lifetime of interconversion. Although all functionals predict the presence of two conformers and that conformer 1 is the most stable, the proportion varies from 67% to 86%. B3LYP/ 6-311+g(d,p) was chosen for use in further calculations due to being to most similar to those results obtained when using mp2.

Appendix II: Input and Relevant Output for Calculation of Fluorescence Emission from the First Excited State

Step 1: Ground state geometry optimisation including frequency calculation with a PCM equilibrium solvation.

Input

```
%chk=step1.chk
%mem=4GB
%nproc=2
#n B3LYP/6-311+g(d,p) opt freq SCRF=(Solvent=ethanol)

Optimisation followed by frequency calculation

0 1
```

Output from Log file: Energy of the ground sate optimized structure in ethanol.

Thermal results from frequency calculation.

```
SCF Done:  E(RM062X) =  -775.998102239   A.U. after    1 cycles
```

```
Zero-point correction=                          0.226836 (Hartree/Particle)
 Thermal correction to Energy=                  0.240154
 Thermal correction to Enthalpy=                0.241098
 Thermal correction to Gibbs Free Energy=       0.186475
 Sum of electronic and zero-point Energies=          -775.771266
 Sum of electronic and thermal Energies=             -775.757949
 Sum of electronic and thermal Enthalpies=           -775.757004
 Sum of electronic and thermal Free Energies=        -775.811628
```

Sum of electronic and thermal Free Energies is equal to G=H-TS

Step 2: TD-DFT single point vertical excitation from the ground state with non-equilibrium linear response water solvation.

Input- The first 6 states have been specified and any coefficients greater than 0.01 be printed.

```
%oldchk=step1.chk
%chk=step2.chk
%mem=4GB
%nproc=2
#n B3LYP /6-311g(d,p) TD(nstates=6,singlet) gfprint pop=(reg,orbital) IOp(9/40=2)
SCRF=(Solvent=ethanol) Geom=Check Guess=Read

qAN4 in ethanol vertical excitation with linear response solvation

0 1
```

Output:

```
Excited State   1:      Singlet-A      3.5146 eV  358.72 nm  f=0.2997  <S**2>=0.000
     55 -> 64         0.01109
     55 -> 65         0.01146
     55 -> 66        -0.01358
     57 -> 64         0.02449
     57 -> 66        -0.01953
     60 -> 63         0.01285
     60 -> 64        -0.06980
     60 -> 65        -0.01902
     61 -> 64         0.05007
     61 -> 65         0.01971
     61 -> 66         0.03618
     62 -> 63         0.69182
     62 -> 64         0.06666
     62 -> 65         0.06430
     62 -> 66        -0.03653
     62 -> 82        -0.01545
     57 <- 64         0.01061
     60 <- 64        -0.01834
     60 <- 65        -0.01014
     61 <- 64         0.01188
     61 <- 66         0.01161
     62 <- 63        -0.04109
 This state for optimization and/or second-order correction.
 Total Energy, E(TD-HF/TD-KS) =  -775.868942587
 Copying the excited state density for this state as the 1-particle RhoCI density.
```

Here absorption would be at 358.72 nm and the oscillator strength of the transition 0.2997. The nature of the transition and the atoms these orbitals correspond to can be found from "Atomic contributions to Alpha molecular orbitals" section in the log file and "Ground to excited state transition densities written to RWF". The main orbitals contributing to this transition are 62 -> 63. This is a pi to pi* transition. 62 is the HOMO with most electron density on C15, N12 and C5 and 63 in the LUMO centred on C24, C15 and N27.

Step 3. First excited state specific approach using ground state equilibrium solvation instead of the excited state equilibrium solvent.

```
%oldchk=step1.chk
%chk=step3.chk
%mem=4GB
%nproc=2
#n M062X /6-311g(d,p) SCRF=(Solvent=ethanol,Read) Geom=Check Guess=Read

qAN4 in ethanol saving the solvent reaction field from ground state.

0 1

NonEq=write

--link1--
%chk=step3.chk
%mem=4GB
%nproc=2
#n M062X /6-311g(d,p) TD(nstates=6,singlet,Root=1) IOp(9/40=2)
SCRF=(Solvent=ethanol,ExternalIteration,Read) Geom=Check Guess=Read

read non-eq solvation from ground state and compute energy of first excited state with state specific
method

0 1

NonEq=read
```

The energy after PCM correction can be subtracted from the ground state energy from step 1 to give the solvent corrected absorption. Here 355.67.

```
After PCM corrections, the energy is -775.863094084 a.u.
```

Step 4. Excited state geometry optimisation with the default linear response solvent.

```
%oldchk=step2.chk
%chk=step4.chk
%mem=4GB
%nproc=2
#n B3LYP/6-311+g(d,p) TD(Read, nstates=6,singlet, Root=1) gfprint pop=(reg,orbital) IOp(9/40=2)
SCRF=(Solvent=water)Geom=Modify Guess=Read Opt=ReadFC

qAN4 in ethanol with freq vertical excitation with linear response solvation

0 1

4 11 10 1 10.0
```

The planar structure is modified to break symmetry before optimisation begins. The output follows the same format as that of Step 2.

```
Excited State   1:      Singlet-A      3.0061 eV  412.45 nm  f=0.4104  <S**2>=0.000
     57 -> 64        -0.01084
     57 -> 66        -0.01764
     60 -> 63         0.01195
     60 -> 64         0.03089
     60 -> 65        -0.01062
     61 -> 64        -0.03678
     61 -> 65         0.01433
     61 -> 66         0.02104
     62 -> 63         0.70107
     62 -> 64         0.01778
     62 -> 65         0.06093
     62 -> 66        -0.01539
     62 -> 82        -0.01401
     60 <- 64         0.01190
     61 <- 64        -0.01324
     62 <- 63        -0.03718
 This state for optimization and/or second-order correction.
 Total Energy, E(TD-HF/TD-KS) =  -775.880046287
 Copying the excited state density for this state as the 1-particle RhoCI density.
```

Step 5. Excited state frequency calculation to confirm the lowest energy structure has been found.

```
%oldchk=step4.chk
%chk=step5.chk
%mem=4GB
%nproc=2
#n B3LYP/6-311+g(d,p) TD(Read, nstates=6,singlet, Root=1) Freq gfprint pop=(reg,orbital)
SCRF=(Solvent=ethanol) Geom=Check Guess=Read

qAN4 in ethanol with excited state freq

0 1
```

Output:

```
                         Item               Value     Threshold  Converged?
 Maximum     Force            0.000010     0.000450     YES
 RMS         Force            0.000002     0.000300     YES
 Maximum     Displacement     0.001063     0.001800     YES
 RMS         Displacement     0.000249     0.001200     YES
 Predicted change in Energy=-5.322396D-09
 Optimization completed.
   -- Stationary point found.
```

To correct emission wavelength the ground state energy must be calculated with the excited state solvent arrangement. This is part of a two-step calculation. In step 6 the excited solvation is saved and in step7 ground state energy is calculated.

Step 6.

```
%oldchk=step5.chk
%chk=step6.chk
%mem=4GB
%nproc=2
#n B3LYP/6-311+g(d,p) TD=(Read,NStates=6,Root=1) Geom=Check Guess=Read
SCRF=(Solvent=Ethanol,ExternalIteration,NonEquilibrium=Save)

qAN4 excited state with state specific solvation

0 1
```

Output

```
After PCM corrections, the energy is  -775.886101169     a.u.
```

Step 7. Ground State energy calculation using solvent calculated in the excited state.

```
%oldchk=step6.chk
%chk=step7.chk
%mem=4GB
%nproc=2
#n B3LYP/6-311+g(d,p) SCRF=(Solvent=Ethanol,NonEquilibrium=Read) Geom=Check Guess=Read

qAN4 ground state calculation with excited state solvation

0 1
```

Output

```
SCF Done:  E(RB3LYP) =  -775.987489213     A.U. after   12 cycles
```

The difference between the energys from Step 6 and 7 can be used to calculate the emission wavelength. From these values this corresponds to a wavelength of 449 nm.

Appendix III: Detailed Listing of qAN Fluorescence Decay Parameters

Globally fitted fluorescence decay parameters of qAN1-4 at an excitation wavelength of 357 nm.

(i) In ethanol

qAN1	τ_1 / ns	A_1	τ_2 / ns	A_2	$\langle\tau\rangle$/ ns	χ^2
420	0.15	0.18	3.15	0.82	2.61	1.08
440	0.15	0.17	3.15	0.83	2.65	1.04
460	0.15	0.18	3.15	0.82	2.60	1.07

Global $\chi^2 = 1.06$

qAN2	τ_1 / ns	A_1	τ_2 / ns	A_2	$\langle\tau\rangle$/ ns	χ^2
430	0.37	0.37	4.82	0.89	4.37	1.13
440	0.37	0.16	4.82	0.84	4.10	1.05
450	0.37	0.16	4.82	0.84	4.13	1.05

Global $\chi^2 = 1.08$

qAN3	τ_1 / ns	A_1	τ_2 / ns	A_2	τ_3 / ns	A_3	$\langle\tau\rangle$/ ns	χ^2
420	0.22	0.21	3.35	0.08	7.97	0.71	6.00	1.12
440	0.22	0.20	3.35	0.10	7.97	0.70	5.96	1.07
460	0.22	0.23	3.35	0.13	7.97	0.64	5.56	1.13

Global $\chi^2 = 1.10$

qAN4	τ_1 / ns	A_1	τ_2 / ns	A_2	$\langle\tau\rangle$/ ns	χ^2
420	0.70	0.05	4.63	0.95	4.45	1.15
440	0.70	0.03	4.63	0.97	4.50	1.07
460	0.70	0.04	4.63	0.96	4.47	1.10

Global $\chi^2 = 1.11$

(ii) In phosphate buffer (pH 7.4)

qAN1	τ_1 / ns	A_1	τ_2 / ns	A_2	τ_3 / ns	A_3	$\langle\tau\rangle$/ ns	χ^2
420	0.07	0.35	2.23	0.01	4.84	0.64	3.14	1.12
440	0.07	0.34	2.23	0.03	4.84	0.63	3.13	1.07
460	0.07	0.32	2.23	0.09	4.84	0.58	3.07	1.07

Global $\chi^2 = 1.087$

qAN2	τ_1 / ns	A_1	τ_2 / ns	A_2	τ_3 / ns	A_3	$<\tau>$*/ ns	χ^2
520	0.31	-0.15	1.48	0.09	4.60	0.75	4.26	1.07
540	0.31	-0.27	1.48	0.07	4.60	0.65	4.28	1.12
560	0.31	-0.30	1.48	0.07	4.60	0.62	4.28	1.12

Global $\chi^2 = 1.13$

*Average of τ_2 and τ_3.

qAN3	τ_1 / ns	A_1	τ_2 / ns	A_2	τ_3 / ns	A_3	$<\tau>$/ ns	χ^2
430	0.23	0.20	4.17	0.14	8.12	0.66	5.96	1.33
450	0.23	0.21	4.17	0.14	8.12	0.64	5.88	1.19
470	0.23	0.24	4.17	0.13	8.12	0.62	5.65	1.25

Global $\chi^2 = 1.26$

qAN4	τ_1 / ns	A_1	τ_2 / ns	A_2	τ_3 / ns	A_3	$<\tau>$/ ns	χ^2
430	0.09	0.25	4.47	0.61	5.27	0.14	3.51	1.02
450	0.09	0.24	4.47	0.60	5.27	0.16	3.53	1.05
470	0.09	0.27	4.47	0.44	5.27	0.28	3.51	1.13

Global $\chi^2 = 1.07$